

## STATUS OF THESIS

Title of Thesis

**Enhancement of Single and Composite Images Based on Contourlet Transform Approach**

I MELKAMU HUNEGNAW ASMARE

hereby allow my thesis to be placed at the information Resource Center (IRC) of the Universiti Teknologi PETRONAS (UTP) with the following conditions.

1. The thesis becomes the property of UTP.
2. The IRC of UTP may make copies of the thesis for academic purposes only.
3. This thesis is classified as

☐

Confidential

☒

Non-confidential

If the thesis is confidential, please state the reason:

\_\_\_\_\_

\_\_\_\_\_

The contents of the thesis will remain confidential for \_\_\_\_\_ years.

Remarks on disclosure:

\_\_\_\_\_

\_\_\_\_\_

Endorsed by



Signature of Author

Permanent : \_\_\_\_\_

Address : \_\_\_\_\_

\_\_\_\_\_

Date: 27/07/09



Signature of Supervisor

Name of Supervisor

Dr. Vijnath S. Asirvadam

Date: 07/07/2009

## UNIVERSITI TEKNOLOGI PETRONAS

### Approval by Supervisor(s)

The undersigned certify that they have read, and recommend to The Postgraduate Studies Programme for acceptance, a thesis entitled "Enhancement of Single and Composite Images Based on Contourlet Transform Approach" submitted by Melkamu Hunegnaw Asmare for the fulfillment of the requirement for the degree of Master of Science in Electrical and Electronic Engineering.

27/07/09

Date

Signature

:

Main Supervisor

:

Date

:

Dr. Vijnath Sagayan Asirvadam  
Senior Lecturer  
Electrical & Electronic Engineering Department  
Universiti Teknologi PETRONAS

Signature

:

Co-Supervisor

:

Date

:

**UNIVERSITI TEKNOLOGI PETRONAS**

**Enhancement of Single and Composite Images Based on  
Contourlet Transform Approach**

**By**

**Melkamu Hunegnaw Asmare**

**A THESIS**

**SUBMITTED TO THE POSTGRADUATE STUDIES PROGRAMME**

**AS A REQUIREMENT FOR**

**THE DEGREE OF MASTER OF SCIENCE**

**IN ELECTRICAL AND ELECTRONIC ENGINEERING**

**BANDAR SERI ISKANDAR,**

**PERAK**

**JULY, 2009**

## DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTP or other institutions.

Signature : Melkamu Honegnaw Asmare

Name : Melkamu Honegnaw Asmare

Date : 27/07/08

## ACKNOWLEDGMENT

First and foremost, all the praise goes to God, who gave me the chance to achieve what I have achieved so far. Second I would like to express my sincere gratitude to my supervisor Dr. Vijanth Sagayan Asirvadam for his guidance and counseling. He has shown me a tremendous patience and care in guiding me through. He has shared with me a lot of insights and experiences, all of which I appreciate. It has been an unforgettable learning experience working with him for the last three semesters. My deepest gratitude goes to my co-supervisor Lila Iznita Izhar for her collaborations.

I am also very grateful to Assoc. Prof. Dr. Syed Abd. Rahman Syed Abu Bakar, Dr. Mohd Yunus Bin Nayan, Prof. P. A. Venkatachalam for serving in my examiners committee and reviewing the thesis. Their valuable and insightful comments are significant for the successful accomplishment of this thesis.

I would like also to thank my brother and mentor Dr. Addisu Hunegnaw, whose constant support and encouragement has brought me here. I am also sincerely grateful to my brother Molla Hunegnaw for all the sympathy and support that you have given me throughout my life.

I take this opportunity to thank all my brothers and sisters who were always available whenever I need them. I always feel lucky for being a member of this wonderful family. I would like to thank all my friends and well wishers for pouring me with their valuable suggestions and standing for me though out the years in the university.

Finally I would like to dedicate this work to my parents, Huneganw Asmare and Bayoush Ayenew, whose pictures do not need any enhancement at all.

Melkamu Hunegnaw Asmare

July, 2009

## **ABSTRACT**

Image enhancement is an imperative step in almost every image processing algorithms. Numerous image enhancement algorithms have been developed for gray scale images despite their absence in many applications lately. This thesis proposes new image enhancement techniques of 8-bit single and composite digital color images. Recently, it has become evident that wavelet transforms are not necessarily best suited for images. Therefore, the enhancement approaches are based on a new ‘true’ two-dimensional transform called contourlet transform. The proposed enhancement techniques discussed in this thesis are developed based on the understanding of the working mechanisms of the new multiresolution property of contourlet transform. This research also investigates the effects of using different color space representations for color image enhancement applications. Based on this investigation an optimal color space is selected for both single image and composite image enhancement approaches. The objective evaluation steps show that the new method of enhancement not only superior to the commonly used transformation method (e.g. wavelet transform) but also to various spatial models (e.g. histogram equalizations). The results found are encouraging and the enhancement algorithms have proved to be more robust and reliable.

## ABSTRAK

Peningkatan mutu paparan sesuatu imej adalah satu perkara yang penting dalam aspek pembinaan perisian untuk pemprosesan imej. Beberapa perisian dalam bentuk algoritma bagi meningkatkan mutu paparan imej *grayscale* atau skala kelabu dihasilkan walaupun aplikasinya kebelakangan ini amat berkurangan. Tesis ini mencadangkan teknik-teknik baru dalam peningkatan mutu paparan imej digital 8-bit berwarna bagi imej-imej tunggal dan komposit. Baru-baru ini, telah dibuktikan teknik *wavelet* tidak semestinya teknik yang terbaik atau *ideal* untuk teknik peningkatan mutu paparan imej. Oleh itu, pendekatan yang baru disyorkan dalam tesis ini bagi meningkatkan mutu paparan imej berdasarkan dengan menggunakan pengubahan dua dimensi yang dipanggil teknik pengubah *contourlet* (*contourlet transform*). Teknik peningkatan mutu paparan imej, melalui *contourlet transform*, yang dicadangkan dalam tesis ini berasaskan mekanisme-mekanisme multiresolution. Penyelidikan dalam tesis ini juga mengaji perwakilan berbeza warna-warna dalam system dapat memberi kesan dalam peningkatan mutu paparan imej yang berwarna. Berdasarkan siasatan ini satu perwakilan warna yang optimum dipilih untuk peningkatan mutu paparan imej-imej tunggal dan komposit. Beberapa penilaian objektif menunjukkan kaedah baru, *contourlet transform*, bukan sahaja kaedah yang lebih berkesan dari proses transformasi selalu digunakan (seperti *wavelet transform*) tetapi lebih baik jika dibandingkan dengan model *spatial* atau konsep ruangan (seperti *histogram equalization*). Keputusan-keputusan penilaian mendapati konsep *contourlet transform* yang dicadangkan dalam tesis ini adalah memuaskan dan menggalakkan dan algoritma peningkatan mutu paparan imej baru ini terbukti lebih robus.

## TABLE OF CONTENSTS

STATUS OF THESIS.....	I
APPROVAL PAGE .....	II
TITLE PAGE .....	III
DECLARATION.....	IV
ACKNOWLEDGEMENT.....	V
ABSTRACT .....	VI
ABSTRAK.....	VII
LIST OF FIGURES .....	XIII
LIST OF TABLES .....	XVIII
CHAPTER ONE: INTRODUCTION.....	1
1.1 Introduction.....	1
1.2 Problem Statement.....	4
1.2.1 Observation.....	5
1.2.2 Mathematical Modeling.....	5
1.2.3 Algorithm.....	5
1.3 Research Objective.....	6



<b>1.4 Scope of the Research.....</b>	<b>7</b>
<b>1.5 Experimental Setup.....</b>	<b>7</b>
<b>1.6 Research Contribution.....</b>	<b>8</b>
<b>1.6.1 Multiresolution Approach.....</b>	<b>8</b>
<b>1.6.2 Single Image Enhancement Approach.....</b>	<b>9</b>
<b>1.6.3 Composite Image Enhancement Approach.....</b>	<b>9</b>
<b>1.6.4 Objective Performance Evaluation.....</b>	<b>9</b>
<b>1.7 Thesis Organization.....</b>	<b>10</b>
 <b>CHAPTER TWO: LITERATURE REVIEW.....</b>	 <b>12</b>
<b>2.1 Introduction.....</b>	<b>12</b>
<b>2.2 Image Contrast Enhancement: Single Image Approaches.....</b>	<b>12</b>
<b>2.3 Composite Image Enhancement Approach.....</b>	<b>21</b>
 <b>CHAPTER THREE: MULTIREOLUTION DECOMPOSITION.....</b>	 <b>27</b>
<b>3.1 Introduction.....</b>	<b>27</b>
<b>3.2 Wavelet Transform.....</b>	<b>28</b>
<b>3.2.1 Multiresolution Analysis and Filter Banks for Wavelet Transform....</b>	<b>31</b>
<b>3.3 Contourlet Transform.....</b>	<b>40</b>

3.3.1 Theory of Contourlets.....	41
3.3.2 Laplacian Pyramid.....	42
3.3.3 Directional Filter Bank Decomposition.....	44
3.3.4 Multiscale and Directional Decomposition (Contourlets) .....	47
3.3.5 Multiresolution Analysis and Filter Banks for Contourlet Transform.....	49
 CHAPTER FOUR: SINGLE IMAGE ENHANCEMENT.....	55
4.1 Introduction.....	55
4.2 Multiresolution Based Image Enhancement.....	56
4.3 Results and Discussion.....	61
4.4 Image Quality Assessment.....	70
4.4.1 The SSIM Index.....	71
4.4.2 Evaluation Setup.....	73
4.4.3 Comparison with the Existing Image Enhancement Technique based on Contourlet Transform .....	78
4.5 Color Space Selection.....	80
4.5.1 Comparison based on Enhancement Performance.....	80

<b>CHAPTER FIVE: COMPOSITE IMAGE ENHANCEMENT.....</b>	<b>84</b>
<b>5.1 Introduction.....</b>	<b>84</b>
<b>5.2 Pixel Based Image Fusion.....</b>	<b>85</b>
5.2.1 MR Analysis.....	87
5.2.2 Color Transformation.....	89
5.2.3 Approximate Coefficients.....	90
5.2.4 Detail Coefficients.....	91
5.2.5 Activity Measure.....	91
5.2.6 Match Measure.....	92
5.2.7 Decision Map.....	93
5.2.8 Combination Map.....	95
5.2.9 MR Synthesis.....	95
<b>5.3 Results and Discussion.....</b>	<b>96</b>
5.3.1 Driving at Night.....	97
5.3.2 Fire Fighting and Rescue Operations.....	98
5.3.3 Multi-Focal Gray Image.....	99
5.3.4 Medical Imaging.....	100
5.3.5 Military Application.....	101
5.3.6 Red Eye Removal.....	102

<b>5.4 Performance Evaluation.....</b>	<b>104</b>
 <b>CHAPTER SIX: CONCLUSION.....</b>	 <b>108</b>
<b>6.1 Introduction.....</b>	<b>108</b>
<b>6.2 Summary of the Thesis.....</b>	<b>109</b>
<b>6.3 Future Works.....</b>	<b>110</b>
<b>LIST OF AWARDS.....</b>	<b>111</b>
<b>LIST OF PUBLICATIONS.....</b>	<b>112</b>
<b>BIBLIOGRAPHY.....</b>	<b>113</b>
<b>APPENDIX A.....</b>	<b>123</b>
<b>APPENDIX B.....</b>	<b>124</b>

## LIST OF FIGURES

Figure 1.1: Fighting Rhinos, Charcoal taken from the two fighting rhinos produced radiocarbon dates of around 31000-32000. (Photo: Time Magazine 13th Feb 1995).....	2
Figure 1.2 Equipment and tools used during the experimental work. (a) Compaq Presario Laptop (b) Matlab 7.0.4 (c) SAMSUNG Digital Camera (d) A4 Convertible Webcam (e) SENSONIC Webcam.....	8
Figure 2.1: Articles found in the IEEE database, with the keyword Image enhance.....	14
Figure 2.2 Image Fusion Venn diagram for two image sensors.....	21
Figure 2.3: The evolution of Multiresolution based Image Fusion research.....	22
Figure 2.4: Articles found in the IEEE database using keyword image fusion.....	23
Figure 3.1 Visualization of Nested Subspaces.....	32
Figure 3.2 Downsampling.....	36
Figure 3.3 Filtering and downsampling for approximate coefficients.....	37
Figure 3.4 Filtering and downsampling for detail coefficients.....	37
Figure 3.5 One stage analysis filter bank.....	37
Figure 3.6 Multiresolution Filter Bank Decomposition.....	38
Figure 3.7 Upsampling.....	38
Figure 3.8 One stage Upsampling.....	38

Figure 3.9 A three stage synthesis filter bank.....	39
Figure 3.10 A one level 2D decomposition wavelet filter bank.....	39
Figure 3.11 Wavelet vs. Contourlets.....	41
Figure 3.12 One level Laplacian Pyramid decomposition.....	43
Figure 3.13 One level Laplacian Pyramid Reconstruction.....	44
Figure 3.14 Directional filter bank frequency partitioning where $l=3$ and there are $2^l = 8$ wedge shaped frequency bands.....	45
Figure 3.15 Two dimensional spectrum partition using quincunx filter banks with fan filters.....	47
Figure 3.16 Application of a rotation operator where the sampling is done by $R_0$ .....	47
Figure 3.17 Contourlet transform framework.....	48
Figure 3.18 a) Contourlet filter bank b) Resulting frequency division.....	48
Figure 3.19 Multiscale nested subspaces generated by the LP.....	50
Figure 3.20 Multiresolution nested subspaces generated by the DFB.....	52
Figure 3.21 Multidirectional nested subspaces generated by the PDFB.....	53
Figure 4.1 Enhanced coefficients versus original coefficients. (a)Velde's method using equation 4.4 with parameters $m = 30$ , $c = 3$ , $s = 0$ , and $p = 0.5$ . (b) Starck's method using equation 4.5 with parameters $m = 30$ , $c = 3$ , $s = 0.6$ , and $p = 0.5$ .....	59
Figure 4.2 Enhanced coefficients versus original coefficients with parameters $a=1$ , $b= 2$ , $\alpha=0.85$ , $c= 4$ .....	60

Figure 4.3 Enhanced coefficients versus original coefficients with parameters $a=1, b=3, c=3, \alpha=0.85$ .....	61
Figure 4.4 The overall general enhancement system.....	62
Figure 4.5 Illustration of the contourlet transform (a) true color image (b) intensity image (c) pyramid transform image (d) contourlet transform image.....	63
Figure 4.6 Example I with enhancement parameters $a=1, b=2, \beta=0.85, \alpha=1, c=3$ .....	63
Figure 4.7 Example II, with enhancement parameters $a=1, b=2, \beta=0.85, \alpha=1, c=3$ .....	64
Figure 4.8 Example III with enhancement parameters $a=1, b=3, \beta=0.85, \alpha=0.7, c=4$ .....	65
Figure 4.9 Example IV with enhancement parameters $a=1, b=1, \beta=0.85, \alpha=0.7, c=4$ .....	66
Figure 4.10 Example V, with enhancement parameters $a=1, b=2, \beta=0.85, \alpha=1, c=3$ .....	66
Figure 4.11 Example VI with enhancement parameters $a=1, b=2, \beta=0.85, \alpha=0.8, c=3$ .....	67
Figure 4.12 Example VII, with enhancement parameters $a=1, b=2, \beta=0.85, \alpha=1, c=4$ .....	68
Figure 4.13 Example VIII, with enhancement parameters $a=1, b=2, \beta=0.85, \alpha=1, c=3$ .....	68
Figure 4.14 Example IX, with enhancement parameters $a=1, b=2, \beta=0.85, \alpha=1, c=3$ .....	69
Figure 4.15 Example X, with enhancement parameters $a=1, b=2, \beta=0.85, \alpha=0.8, c=3$ .....	70
Figure 4.16 Objective evaluation test image (a) Perfect reference image (b) test image (c) HE Method (d) CLAHE (e) wavelet enhanced image (f) contourlet enhanced image.....	75
Figure 4.17 RGB profile of the result from the Figure 4.16 (a) Profile of the perfect image	

(b) Profile of the test image (c) profile of the HE Method	
(d) profile of the CLAHE Method (e) Profile of the Wavelet Method	
(f) Profile of the Contourlet Method.....	76
Figure 4.18 Comparison of all the four methods for all 31 images.....	77
Figure 4.19 Comparison of our method with the P. Feng contrast enhancement method	
(a) True color retinal image (b) green band (c) our method (d) P. Feng method.....	79
Figure 4.20 Enhancement of the Chair image in different color spaces	
based on the contourlet transform method.....	82
Figure 5.1 Generalized pixel based MR image fusion scheme with two input	
images and a composite image output.....	87
Figure 5.2 Sample Images a) visual image b) night vision image.....	88
Figure 5.3 Contourlet transform coefficients of the source images, a) visual image	
b) night vision image.....	89
Figure 5.4 HSI decomposition of the visual image a) hue b) saturation c) intensity.....	90
Figure 5.5 Visualization of the Activity measures (a) Visual Image (b) night vision image.....	92
Figure 5.6 Visualization of the match measure.....	93
Figure 5.7 Decision Map (a) Visual Image (b) night vision image.....	95
Figure 5.8 Composite Image.....	96
Figure 5.9 Night time driving (a) Visual camera (b) night vision camera	
(c) composite image.....	97
Figure 5.10 Fire fighting and rescue operation (a) visual camera (b) infrared camera	



(c) composite image.....	98
Figure 5.11 Multi-focal gray scale image (a) Near focused (b) far focused	
(c) composite image.....	99
Figure 5.12 Medical imaging (a) MRI image (b) CT image (c) composite image.....	100
Figure 5.13 Military application (a) Infrared image (b) visual image (c) fused image.....	101
Figure 5.14 Red eye removal (a) Image taken when the flash is on	
(b) image taken when the flash is off (c) fused image.....	103
Figure 5.15 Objective evaluation test image (a) original reference image	
(b) right side blurred image (c) left side blurred image (d) contourlet method	
(e) wavelet method (f) simple averaging.....	105
Figure 5.17 Objective evaluation test image (a) original reference image	
(b) right side blurred image (c) left side blurred image	
(d) contourlet method (e) wavelet method (f) simple averaging.....	107
Figure B.1 GUI for single image.....	124
Figure B.2 GUI for composite image.....	125

## LIST OF TABLES

Table 4.1 Objective evaluation result for the example image from Figure 4.16.....	74
Table 4.2 Objective Evaluation for 31 images.....	77
Table 4.3 Computational time for retinal images.....	80
Table 4.4 ASSIM for the test images using in three different enhancement algorithms.....	81
Table 5.1 Performance evaluation result of the test image on Figure 5.15.....	104
Table 5.2 Performance evaluation result of the test image on Figure 5.16.....	106

## **CHAPTER ONE: INTRODUCTION**

### **1.1 Introduction**

In early stages of human development, human beings tend to communicate through symbols and drawings. The emergence of cave art about 30,000 years ago is widely believed to be the evidence that this time they had developed sophisticated capacities for symbolization and communication [N. Humphrey, 1998]. The cave painting contains hundreds of different paintings most of which are animals (e.g. Fighting Rhinos shown in Figure 1.1) which tells that humans started using imagery very long ago.

In the industrial revolution, there have been a lot of paradigm shifts in imaging and photography. The medium of image communication has changed from simple cave and wall paintings to images and videos captured by cameras, fax machines, and televisions. Recently in the information technology era the advancement of digital imaging leads to images that can be captured and transmitted through digital mediums like computers and cell phones.

At present, human beings are extensively using images. Cameras have become integral parts of the modern day life. It is familiar to see cameras installed in banks, traffic intersections, libraries, shopping malls and etc for continuous surveillance. It is also customary to find cameras in every person's pocket to record different moments. Even cameras are miniaturized and included in other electronic devices as an additional accessory. Usually they are integrated in cell phones, laptops, palmtops and automobiles.

Advances in sensor technology, microelectronics and digital signal processing, have produced cameras which are capable to overcome a lot of problems. Blur resulting from the speed of the camera, the aperture and the optical lens has been properly solved.

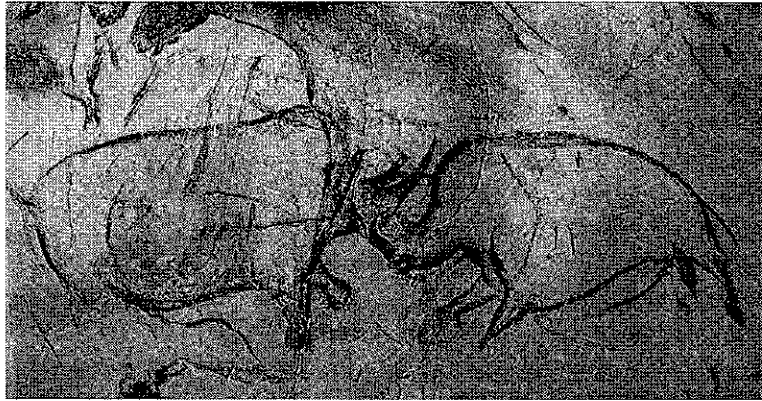


Figure 1.1: Fighting Rhinos, Charcoal taken from the two fighting rhinos produced radiocarbon dates of around 31000-32000 (Photo: Time Magazine 13th Feb 1995)

Most cameras have good color filtering arrays so it is now possible to overcome color artifacts which further degrade the quality of the images. Even some kind of intelligence has been incorporated into cameras. A number of camera models have eye location detector and face recognition capability.

However, there is an inherent problem of all visual cameras. All need a threshold of ambient illumination to take a good picture. In the absence of such illumination, images taken by visual cameras have dark shadows and low contrast. We may also end up with having low quality digital images caused by high dynamic range, noise, poor contrast and non-uniform illumination. So far this problem has been unavoidable. Here we propose a single image enhancement approach to alleviate this problem. However the single image approach is highly dependent on the quality of the image sensor by which the image is taken. Either by design or because of observational constraints, most image sensors are specialized in capturing only limited aspects of the image. To solve this problem we also propose an image enhancement technique based on composite images. Important features from two or more images will be combined to produce a single more complete enhanced image.

These image processing techniques are highly dependent on the digital representation of the input images. Non linear approximation (NLA) techniques have so far proved to be very important in this aspect [Ronald A., 1998]. The wavelet transform is one of the well known NLA methods. It is now accepted to be the ultimate choice for image processing [M. Vetterli, 2001, K. Amolins et al 2007]. But the application of wavelet transform in two dimensional (2D) digital images poses problems. This is because the wavelet transform in 2D system is actually defined as tensor products of the one dimensional (1D) wavelets. One dimensional wavelets are good in capturing point discontinuities which are usually considered as a zero dimension. In reality natural image discontinuities are most of the time 1D, in which wavelet usually fails to capture [M. N. Do, 2001].

Observing this problem several researchers have proposed better NLA systems in 2D. One of these techniques is the contourlet transform which is referred to as a true 2D NLA technique. So far the contourlet transform has proved to have a better denoising and compression capability than the conventional wavelet transform [M. N. Do, M. Vetterli, 2005]. It is sparser as compared to wavelet transform. It easily captures the smoothness in the natural images which has a decisive effect in the human visual system. In this thesis we employ the rich features of the contourlet transform in image enhancement for both single and composite image approaches.

The contourlet transform is actually an extension of the wavelet transform in 2D using separable and directional filter banks. It is composed of basis functions which are oriented in varying directions in multiple scales with flexible aspect ratios. With this rich set of basis functions it can successfully capture the edge information in digital natural images with small number of coefficients. Therefore, the contourlet transform is more sparse and suited for processing digital images.

One possible digital image processing approach is to employ the contourlet transform in applications of image contrast enhancement and dynamic range compression algorithms on the distorted single images. The other approach is to enhance image features by using

the information gathered from multiple images from different sensors. For example, we can combine an image from a night vision camera with an image from a visual camera. In this case, the night vision camera is capable of taking images in low light condition but it cannot capture any color information. On the other hand, the visual camera can take the color information but the image captured will have low contrast and dark shadows. Combining these two images can successfully capture all the relevant information including color. This process is usually called image fusion. Since the fused image generally possesses more scene information than any single input image, image fusion can be considered as an image enhancement process. The multi sensor image fusion has become less expensive as the price of image sensors has dropped in the last decade.

From real life point of view, consider the human visual system, human beings have a superior vision in a good lighting condition. But at night or in low lighting condition our vision is very poor as compared to other animals like cats. On the other hand, human beings are forced to work more and more at night. There is an urgent need to perform well at our jobs in low lighting conditions. Thus one possible solution would be to apply a composite image enhancement approach.

## **1.2 Problem Statement**

We are specifically interested in the study of image processing approaches for improving the visibility of low quality digital images. Image enhancement for both single and composite image approaches can be done in several ways. Among the various frameworks in which it has been formulated, the multiresolution approach is one of the most intensively studied and used in practice.

In this thesis we reframe multiresolution based image enhancement and fusion systems. Although scene understanding is a difficult task; we have tried to solve some basic problems. We have characterized the problem statement in three different points of view,

namely: observation point of view, mathematical modeling point of view, and algorithm point of view.

### **1.2.1 Observation**

When we observe the human visual system, it is not perfect especially in low lighting condition. Therefore it must be improved that means we need a vision assistant. In this case the camera will be an avoidable tool. However either by design or observational constraints the images captured by cameras are not also always perfect and complete, thus image enhancement should be done before using them for further processing or for visual perception.

### **1.2.2 Mathematical Modeling**

In engineering design problems it is a common practice to mathematically model a physical problem. The mathematical model should always be efficient and as much as possible must successfully approximate the physical problem at hand. So far one of the NLA methods namely the multiresolution (MR) is the best model for image processing tasks which involve image enhancement. Most of the existing MR algorithms use the wavelet transform. But wavelets are less suited to deal along smooth contours in an image. Due to the limitation of wavelet transform in capturing the intrinsic geometrical structures of image's key information; a contourlet transform technique which is more robust and efficient in doing so is proposed in this study.

### **1.2.3 Algorithm**

The need for fast computation in a possible extension of our work which is a high frame rate video enhancement application motivates our investigation into computationally simple algorithms. Thus in both the algorithms a computational friendly approach is followed.

### 1.3 Research Objective

The main goal of this thesis is to develop an innovative color image enhancement technique for improving the visibility of low quality digital images caused by high dynamic range, noise, poor contrast and low or non uniform illumination. The system can be used for security surveillance, driver's assistance at night or in low light condition, fire fighting operation, military and police target detection and acquisition, medical imaging, multimedia systems and search and rescue operations, etc.

We aim to solve some of the underlying problems using two approaches: a single image enhancement approach and composite image enhancement approach. The NLA systems using filter banks of wavelet and contourlet transform are studied thoroughly. The contourlet transform is the major NLA method used in this work but we will also employ the wavelet transform for comparison purposes. In both enhancement approaches; we employ the contourlet transform to generate a system which can meet the following major objectives:

1. Should not discard any salient information contained in the input image/images.
2. Should not introduce any artificial artifacts or inconsistencies which can mislead a human observer or any subsequent image processing steps.
3. Must be reliable, robust and as much possible tolerant to imperfections such as noise.

### 1.4 Scope of the Research

The proposed system is applied for enhancing low quality natural still images which are taken in different conditions. First a single image enhancement approach is being investigated. The performance of the single image approach is usually limited to the performance of the sensor in which the image is taken. For instance in cases of visual



cameras, images taken in a very poor illumination condition are always covered by dark shadows which is impossible to remove. This time we should capture this lost information by another camera which is capable of capturing the lost parts. Thus a composite image enhancement approach which increase the performance of the overall image enhancement process. These algorithms are implemented on both wavelet and contourlet transform methods. In the composite image enhancement system, a pixel based fusion algorithm is developed and tested. The input images for the fusion system are assumed to be pre registered/aligned using existing simple image alignment techniques.

### **1.5 Experimental Setup**

A software package has been developed based on the proposed enhancement algorithm. This package is implemented in MATLAB7.0.4 development environment in windows XP operating system platform. The MATLAB image processing toolbox is used for basic mathematical operations which involve images. The software packages supports JPEG, BMP, PNG, and TIFF image file formats.

The software is developed in Compaq Presario V3000 Notebook, with AMD processor, with processing speed of 2GHz and 2GB DDR SDRAM memory. The digital cameras used to capture the experimental test images include SAMSUNG S630 Digital camera with SAMSUNG SHD Lens, A4 TECH USB Convertible PC Camera which has a light sensor therefore it is capable of capturing both night vision images and visual images depending on the ambient illumination. SENSONIC Webcam is also used to capture some of the visual images. Figure 1.2 shows all the equipments and tools used in the experiment.

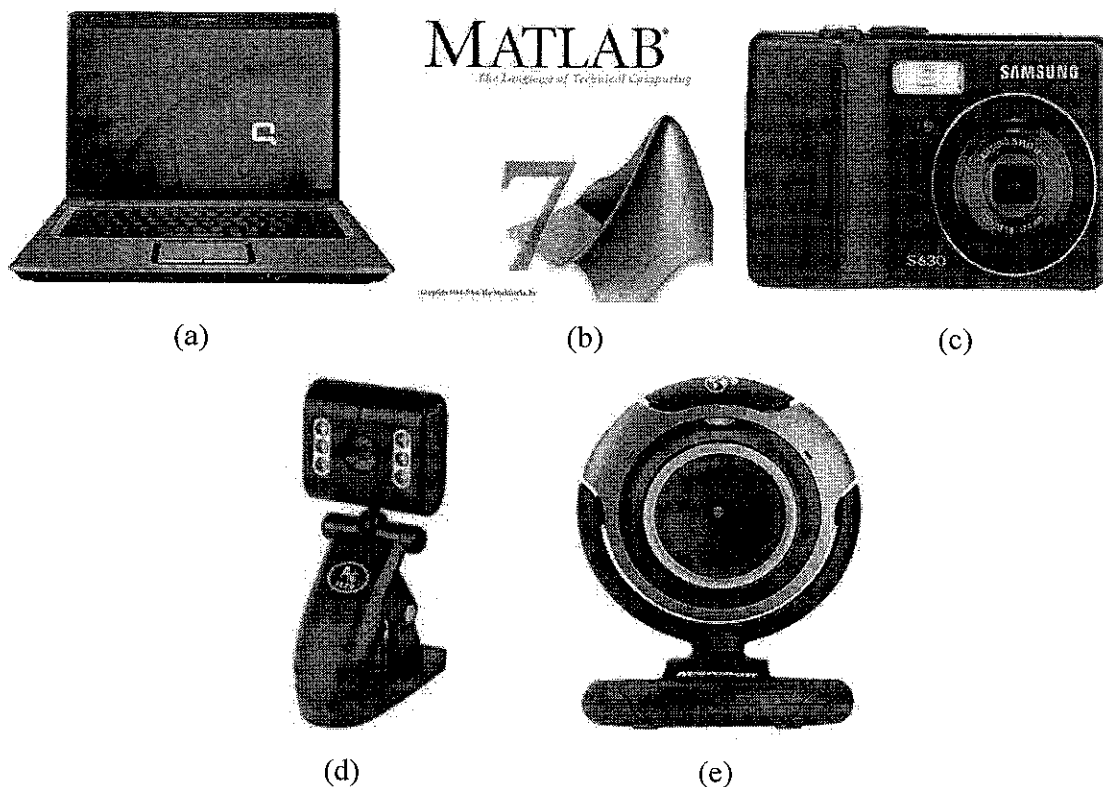


Figure 1.2 Equipment and tools used during the experimental work. (a) Compaq Presario Laptop (b) Matlab 7.0.4 (c) SAMSUNG Digital Camera (d) A4 Convertible Webcam (e) SENSONIC Webcam

## 1.6 Research Contribution

The main contributions of the thesis look into the application of multiresolution approach called ‘contourlet transform’ for image enhancement. The following are list of the main contributions of this work

### 1.6.1 Multiresolution Approach

A newly developed mutliresolution approach, called contourlet transform is used in this thesis for decomposing images into multiple resolutions. The contourlet transform is composed of basis images oriented at varying directions in multiple scales with flexible

aspect ratios. With this rich set of basis images, the contourlet transform has effectively captured the one dimensional smooth contours, edges and lines of the image which are dominant features of all natural images with only few numbers of coefficients. Since it can efficiently and successfully approximate the image the resultant enhanced image is better as compared to the wavelet transform.

### **1.6.2 Single Image Enhancement Approach**

For single image enhancement approach a new enhancement function is proposed. This function meets all the requirements of transfer functions in multiresolution image enhancement applications. It has proved to be efficient, more robust and does not introduce artifacts.

### **1.6.3 Composite Image Enhancement Approach**

This approach is implemented in contourlet transform domain. The combination is based on the works of Burt and Kolczynski. But unlike their method, in this thesis a different fusion rule is proposed for detail and approximate coefficients. If the decomposition level is high enough, approximate and detail coefficients carry different information. The approximate coefficients are combined by simple averaging while the detail coefficients pass through the match and activity measures. Doing so has reduced the noise and it also decreased the computational complexity.

### **1.6.4 Objective Performance Evaluation**

The structural similarity index measure proposed by Zhou Wang et al is used as a landmark. Their method compares the whole image at once. However image signals are non-stationary thus it is a good practice to measure over local regions rather than measuring through the entire image. In this thesis a sliding window ( $w$ ) approach is proposed which gives a better objective evaluation.

### **1.7 Thesis Organization**

The thesis is organized as follows; chapter two presents a literature survey of the conventional color image enhancement and image fusion methods. It discusses some of the most related algorithms to our work. Some spatial domain image contrast enhancement algorithms are discussed here. Image enhancement algorithms based on multiresolution approaches like the wavelet transform and curvelet transform are also investigated. This chapter also presents some of the existing works related to multiresolution based image fusion techniques.

In Chapter Three, multiresolution representation approaches are briefly discussed. Two dimensional wavelet transform decomposition is first studied, then the filter bank representation of the 2D wavelet decomposition is presented. The theory of contourlet transform is also presented in this chapter. Image decomposition using contourlet transform is discussed and formulated with the filter bank representation.

Chapter four proposes multiresolution based color image enhancement algorithm using a single image. A technique for improving the visibility of low quality digital images caused by high dynamic range, noise, poor contrast and very low illumination is developed. A new non linear transformation function is proposed to elevate the values of low intensity pixels. A color space selection method based on the performance of enhancement algorithms when they are applied in different color spaces is also done. Quantitative evaluation methods are used to measure the performance of the enhancement algorithm.

Composite image enhancement approach (Image fusion) is studied in chapter five. An image fusion system which is called fusion by selection of appropriate contourlet coefficients, which works well for both multi modal and multi focal images is developed. The fusion algorithm is successfully applied for combining several multi-modal and

multi-focal images. The performance of the system is studied for different case study images.

Finally chapter six gives concluding remarks and a future outlook for possible research direction in the field of image enhancement and fusion systems.

## **CHAPTER TWO: LITERATURE REVIEW**

### **2.1 Introduction**

Image processing is a field of digital signal processing that uses an image, such as photographs or frames of video as input; and the output can be either an image or a set of characteristics or parameters related to the image. It is the fastest growing field of digital signal processing. Several factors promise the growth of the field. The major one is the declining cost of computers and required equipments. The other factors which indicate the continued growth of the field are parallel processing, efficient charge coupled devices (CCD), new memory technologies for large data storages, and high resolution color display systems etc.

Image processing can be applied in different areas. Medical imaging, multimedia systems, computer vision, remote sensing, security surveillance, process monitoring and inspection, oil and gas exploration are among the vast areas of application. A huge collection of literatures and research papers can be found in a number of image processing applications. Among these enormous collections of literature we are mainly interested in areas of multiresolution based color image contrast enhancement (single image approach) and multiresolution based image fusion methods (composite image approaches). Literature review of some of the existing methods which are more relevant to our focus is given below.

### **2.2 Image Contrast Enhancement: Single Image Approaches**

Image contrast is the variation in the intensity of an image formed by an optical system as black and white bars. It can also be defined as the difference in brightness between two adjacent pixels. We call the contrast of any (small) part of an image the local contrast. The global contrast is defined as the average local contrast of smaller image fractions. An

image with a high global contrast causes a global feeling of a detailed and variation-rich image. As opposed to it, an image with a lower global contrast contains less information, fewer details, and appears more uniform.

Image contrast enhancement is an image processing branch aimed at assisting image visual analysis. It is widely used in medical [J. Tang et al 2007, P. Feng et al 2007, Etta D. et al 2000, Monica A. et al 2006], biological [J. Singhai et al, 2007] and multimedia systems [D. Xiao et al 2007, P. Dong-Liang et al 2005, K. Chen et al 1992, G. Petsching et al 2004, Zhengyou Wang et al 2006] to improve the image quality. Theoretically, image enhancement methods may be regarded as an extension of image restoration methods.

In image enhancement research society, it is a well known fact that image enhancement has no general unifying theory. This is because the enhancement process is problem oriented. Thus an algorithm that works for some types of images may not work with other images. Based on the motivation of development, enhancement algorithms found in the literature can be general algorithms [G. Apostolopoulos et al 2007, Iyad Fayez 2008] proposed to be utilized by other image applications, or algorithms developed for specific applications or types of images such as X-rays [Shojiro Maki et al 2006] and mammograms [J. Tang et al 2007, Etta D. et al 2000].

Image types can be broadly classified as binary images, indexed images, gray scale images and true color images. But for this thesis we mainly concentrate on gray scale and true color images [R. C. Gonzalez, R. E. Woods, 1992]. A gray scale image is a data matrix whose values represent intensities within some range. The computer represents a gray scale image as a matrix where each element of the matrix represents the value of one pixel. A true color image is an image in which each pixel is specified by three values: one each for the red, blue, and green components of the pixel's color. The computer stores true color images as m-by-n-by-3 data array that defines red, green, and blue color components for each individual pixel. In the past most image processing algorithms are

developed for gray scale images. But nowadays gray images are almost obsolete especially in multimedia applications [D. Xiao, Jun Ohya, 2007]. Despite its abundance in all the systems, color image enhancement is not a well explored area, although demands of it are very high in recent years in all multimedia applications. In this thesis we mainly focus on color images.

Images in the computer system are represented as arrays of numbers. Gray scale images are represented by one dimensional array while color images are represented by multidimensional arrays usually three dimensional. It is a common practice to apply gray scale algorithms to each band of color images with a slight modification. Thus the computational complexity for color image enhancement can simply be considered as three fold. We can find tremendous amounts of enhancement algorithms in the literature. The graph in Figure 2.1 shows the number of articles found in the past from the IEEE database using a general search by the keyword “Image Enhancement” [<http://ieeexplore.ieee.org/Xplore/dynhome.jsp>, accessed on Jan. 29, 2009]. It illustrates the increasing interest of the research in image enhancement. In the following paragraphs we will present the survey of some of the most common image enhancement algorithms.

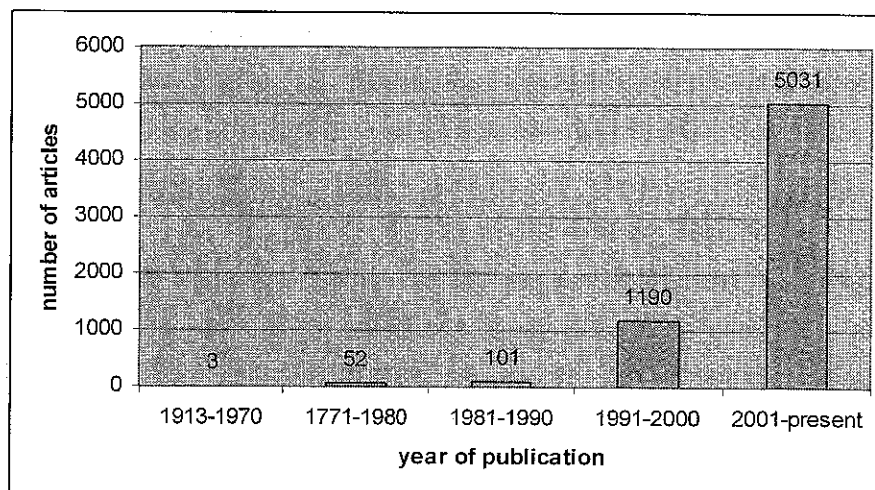


Figure 2.1: Articles found in the IEEE database, with the keyword Image enhancement



Histogram equalization (HE) [Hummel R. et al, 1977] is one of the most well known methods for contrast enhancement. It involves transforming the intensity values so that the histogram of the output image approximately matches a specified histogram. HE works well for images which do have either very dark or very bright scenes only. It fails to perform well for scenes which contain both bright and dark scenes at the same time. This approach is generally useful for images with poor intensity distribution. It has also a noise amplification effect for noisy images.

Contrast limited adaptive histogram equalization (CLAHE) [Pizer S.M. et al, 1987] is an improved version of HE. While HE works on the entire image CLAHE operates on small regions in the image called tiles. Each tile's contrast is enhanced so that the histogram of the output region approximately matches a specified histogram. After performing the equalization, CLAHE combines neighboring tiles using bilinear interpolation to eliminate artificially induced boundaries. The CLAHE algorithm is computational intensive as interpolation is used.

The HE and CLAHE are not commonly used for processing color images because of their strong contrast enhancement which leads to excessive noise or artifacts and cause the image to look unnatural. However, they can be applied to a color image by applying them to each spectral band of color. But doing so changes the color information of the image thus the image must first be converted to another color space which separates the color information from the intensity and apply the enhancement on the intensity component only.

Several varieties of HE exist in the literature. Contrast enhancement by using multi-scale adaptive histogram equalization (MAHE) proposed by [Yinpeng Jina et al, 2001] is one of them. It introduces an approach which utilizes multi-scale analysis. They used the quadratic spline wavelet function, which has compact support and is continuously differentiable to form the multiscales. By doing so they claim that they can selectively enhance features of interest by modifying corresponding components in the transform

domain. Their final result shows a remarkable improvement in applications for computer tomography images.

There exist other methods which directly enhance each band of the color image. The Retinex concept which was introduced by [Edwin H. Land, 1983] as a model for human color vision constancy is one of them. The single scale Retinex (SSR) method consists of applying the following transform to  $k^{\text{th}}$  band of the color image:

$$R_k(x, y) = \log(I_k(x, y)) - \log(F(x, y) * I_k(x, y)) \quad (2.1)$$

where  $R_k(x, y)$  is the Retinex output,  $I_k(x, y)$  is the image distribution in the  $k^{\text{th}}$  spectral band,  $F(x, y)$  is a Gaussian function, and  $*$  is convolution.

A gain/offset is applied to the Retinex output which clips the highest and lowest signal excursions. The Retinex method is efficient for dynamic range compression, but does not provide good tonal rendition.

The Multiscale Retinex (MSR) [Z. Rahman et al, 1996] combines several SSR outputs to produce a single output image which has both good dynamic range compression and color constancy. Color constancy may be defined as the independence of the perceived color from the color of the light source and good tonal rendition. The MSR can be defined as:

$$R_{MSR_k} = \sum_{j=1}^N w_j R_{k,j} \quad (2.2)$$

$$R_{k,j}(x, y) = \log(I_k(x, y)) - \log(F_j(x, y) * I_{k,j}(x, y))$$

$N$  is the number of scales,  $R_{k,j}$  is the  $k^{\text{th}}$  spectral component of scale  $j$  of the MSR output, and  $w_j$  is the weight associated with the scale  $j$ . The Gaussian function  $F_j$  is given by:

$$F_j(x, y) = Ke^{-\frac{r^2}{c_j^2}} \quad (2.3)$$

where  $c_j$  defines the width of the Gaussian function.

MSR generally works well for various types of images. However, it has some drawbacks. Since it needs to process all the spectral bands of the color image, it takes a long time to enhance images. Thus it is not possible to use this method in real time applications and for high frame rate video images.

The MSR is the first to introduce the concept of multiresolution for contrast enhancement. It performs dynamic range compression and can be used for different image processing goals. Improvements of the algorithm have been presented in [K. Barnard and B. Funt, 1999], leading to better color fidelity.

MSR softens the strongest edges and keeps the faint edges almost untouched. The opposite approach was proposed by [Koen Vande Velde, 1999] using the wavelet transform for enhancing the faintest edges and keeping untouched the strongest. The strategies are different, but both methods allow the user to see details which were hardly distinguishable in the original image, by reducing the ratio of strong features to faint features.

The Velde's wavelet approach consists of first transforming the image using the dyadic wavelet transform (two directions per scale). Then the wavelet coefficients,  $x$ , at scale  $j$  and at pixel position  $k$  are multiplied by  $y(x)$ , where  $y(x)$  is defined by:

$$y(x) = \begin{cases} \left(\frac{m}{c}\right)^p & \text{if } |x| < c \\ \left(\frac{m}{x}\right)^p & \text{if } c \leq |x| < m \\ 1 & \text{if } |x| \geq m \end{cases} \quad (2.4)$$

The parameters  $c$ ,  $m$  and  $p$  are used to control the nonlinearity of the enhancement function. Finally the enhanced image can be found by taking the inverse wavelet transform of the modified coefficients. This approach has a lower computational time and the results obtained are good.

Another paper proposed by [Xuli Zong et al, 1996] presents an approach which addresses both de-noising and contrast enhancement using the multiresolution approach in the discrete wavelet transform (DWT) domain. In this paper, they investigated methods of hard thresholding and soft thresholding wavelet shrinkage for noise reduction. An advantage of soft thresholding is smoothness, while hard thresholding preserves features. They showed that two conflicting objectives: de-noising and contrast enhancement can be accomplished simultaneously.

The soft thresholding is formulated as:

$$u(x) = T(v(x), t) = \text{sign}(v(x))(|v(x) - t|) \quad (2.5)$$

where threshold  $t$  is directly proportional with the noise level and  $u(x)$  is the result of threshold and has the same sign as the coefficient  $v(x)$ .

DWT coefficients are processed for noise reduction by the above thresholding formula. For feature emphasis nonlinear processing is generalized in this paper to incorporate hard thresholding to avoid amplifying noise and remove small noise perturbations within

middle scales of analysis. The generalized adaptive gain (GAG) nonlinear operator is defined by:

$$E_{GAG}(v) = \begin{cases} 0 & |v| < T_1 \\ \text{sign}(v)T_2 + \bar{a}(\text{sigm}(c(u-b)) - \text{sigm}(-c(u+b))) & T_2 \leq |v| \leq T_3 \\ v & \text{otherwise} \end{cases} \quad (2.6)$$

where  $v \in [-1, 1]$ ,  $\bar{a} = a(T_3 - T_2)$ ,  $u = \text{sign}(v) \frac{|v| - T_2}{T_2 - T_3}$ ,  $b \in (0, 1)$ ,  $0 \leq T_1 \leq T_2 < T_3 \leq 1$ ,  $c$  is

a gain factor and  $a$  can be computed using equation 2.7 and 2.8:

$$a = \frac{1}{\text{sigm}(c(1-b)) - \text{sigm}(-c(1+b))} \quad (2.7)$$

$$\text{sigm}(v) = \frac{1}{1 + e^{-v}} \quad (2.8)$$

Several multiresolution image enhancement methods based on the wavelet transform exist in the literature. [J. Qin et al 2006, A.F. Laine et al 1994, and F. Sattar et al 1997] are some examples. However, wavelet bases present some limitations, because they are not well adapted to the detection of highly anisotropic elements, such as alignments in an image, or sheets in a cube. Recently, other multiscale enhancement algorithms have been developed, which use in particular curvelets [Jean-Luc Starck et al, 2003], and contourlets [G. Cheng et al 2008, E. Nezhadraya et al 2006, P. Feng 2007] and these are very different from wavelet-like systems. They take the form of basis elements which exhibit very high directional sensitivity and are highly anisotropic.

The enhancement approach which is proposed by [Jean-Luc Starck et al, 2003] is a modified version of the Velde's algorithm. It is a multiresolution approach in the curvelet transform domain. The argument they used is the approach successfully removes the Gaussian noise. As the noises are not parts of structural information of the image, the

curvelet transform will not generate coefficients for the noise. They introduced a noise standard deviation  $\sigma$  in  $y(x)$ .

$$y(x) = \begin{cases} 1 & \text{if } x < c\sigma \\ \frac{x - c\sigma}{c\sigma} \left( \frac{m}{c\sigma} \right)^p + \frac{2c\sigma - x}{c\sigma} & \text{if } x < 2c\sigma \\ \left( \frac{m}{x} \right)^p & \text{if } 2c\sigma \leq x < m \\ \left( \frac{m}{x} \right)^s & \text{if } x \geq m \end{cases} \quad (2.9)$$

The parameters  $p$ ,  $m$ ,  $s$  and  $\sigma$ , need to be defined and enhancement algorithm has a greater tendency of detecting edges corrupted by noise.

They claimed that the curvelet method [Jean-Luc Starck et al, 2003] is superior to other multiresolution methods in enhancing noisy images. But the implementation of the curvelet transform is very difficult because it is defined in the polar coordinates. Thus converting back to the Cartesian coordinate is a difficult task and therefore it is computationally intensive.

By spotting this problem [P. Feng et al, 2007] did the same work like Jean-Luc Starck's method, using the same enhancement function and similar procedures, for retinal image application but by using the contourlet transform whose construction is simpler as compared to the curvelet transform. They claimed that applying the contourlet transform can help them to easily detect retinal lesions and highly helpful for diagnosis.

Detail explanation about the contourlet transform is presented in chapter three. A lot of algorithms are found in the literatures which employ the fuzzy theory in image enhancement application. But Fuzzy logic is not our main focus so we did not discuss them here. [M. Hamnandlu et al 2003, D. B. Handari et al 1993, E. Nezhederya et al 2006] are methods which use the fuzzy theory, interested readers can refer to them. We

propose a new generic enhancement technique applying multiresolution by the contourlet transform. We also propose a new transformation function which is capable of reducing noise while increasing the contrast. A more detail explanation of this technique is given in chapter four.

### 2.3 Composite Image Enhancement Approach

The composite image enhancement approach is a process of combining relevant information from two or more images into a single image. The resulting image will be more informative than any of the input images. Since the fused image generally possesses more scene information than any single input image, this method can also be considered as an image enhancement process. Generally image fusion has two crucial advantages. It improves reliability by redundant information and improves capability by complimentary information. Figure 2.2 shows the Venn diagram representation for two sensors.

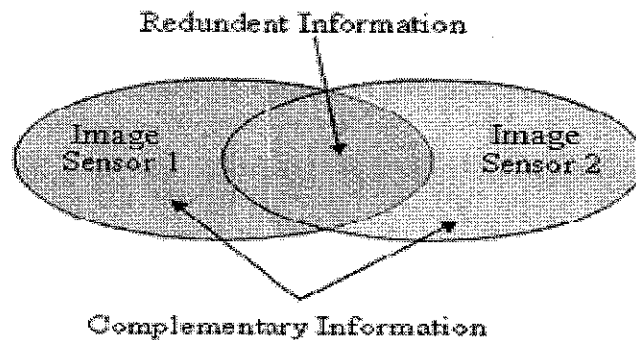


Figure 2.2 Image Fusion Venn diagram for two image sensors

The primitive fusion schemes perform the fusion right on the source images, which often have serious side effects such as reducing the contrast. Other methods first transform images to a more stable form of representation and the fusion is done on the transformed coefficients. There are different combination methods which are working in the existing fusion algorithms. These include: Heuristic, Logic OR/AND, Averaging, Least square, Geometric mean, Bayesian, Possibility (Fuzzy). Our proposed approach is called

Combination by Selection. The other methods have no selection paradigm. We thought that incorporating the informed selection capability will produce a better result.

When Pyramid transform was introduced three decades ago, some sophisticated approaches began to emerge. Researchers found that it would be better to perform the fusion in the transform domain. Pyramid transform appears to be very useful for this purpose. The basic idea is to construct the pyramid transform of the fused image from the pyramid transform of the source images, and then the fused image is obtained by taking the inverse pyramid transform. Pyramid transform can provide information on the sharp contrast changes, and human visual system is specifically sensitive to these sharp contrast changes. It can also provide both spatial and frequency domain localization.

Since then, image fusion began to receive increasing attention. Figure 2.3 shows the evolution of image fusion methods. The graph in Figure 2.4 shows the number of relevant articles found in recent years using a general search by the key word “Image Fusion” from IEEE database [<http://ieeexplore.ieee.org/Xplore/dynhome.jsp>, accessed on Jan. 29, 2009].

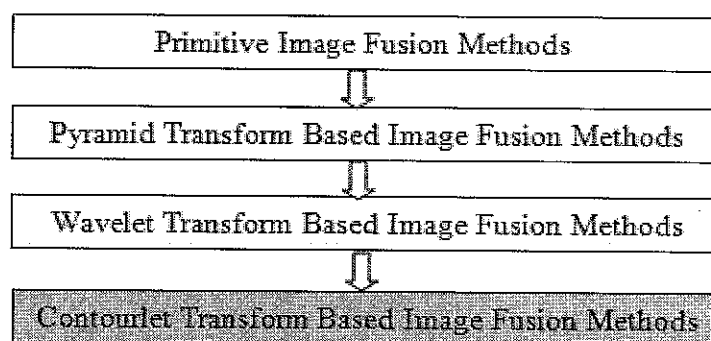


Figure 2.3: The Evolution of Multiresolution based Image Fusion research

Image fusion is an intrinsic behavior of all animals. For instance, human beings use two eyes i.e. two visual sensors to form one complete composite image. Doing so we are capable of having a wider view with no blurring and we can accurately estimate the location and velocity of the objects in the scene. Another very fascinating example is the



visual systems of rattlesnakes. These snakes have two eyes (visual sensors) and an infrared sensor. The image formed from these two different sensors is combined into one image by bi-modal neurons and form a complete and detailed image. Considering this analogy several researchers have tried to perform image fusion using intelligent (neural) network approaches [H. Singh et al, 2004]. In this thesis we mainly focus on multiresolution based image fusion. The review of some of multiresolution based image fusion algorithms is given below. The number of literatures surveyed here is by no means exhaustive but we mainly focus on the major papers available.

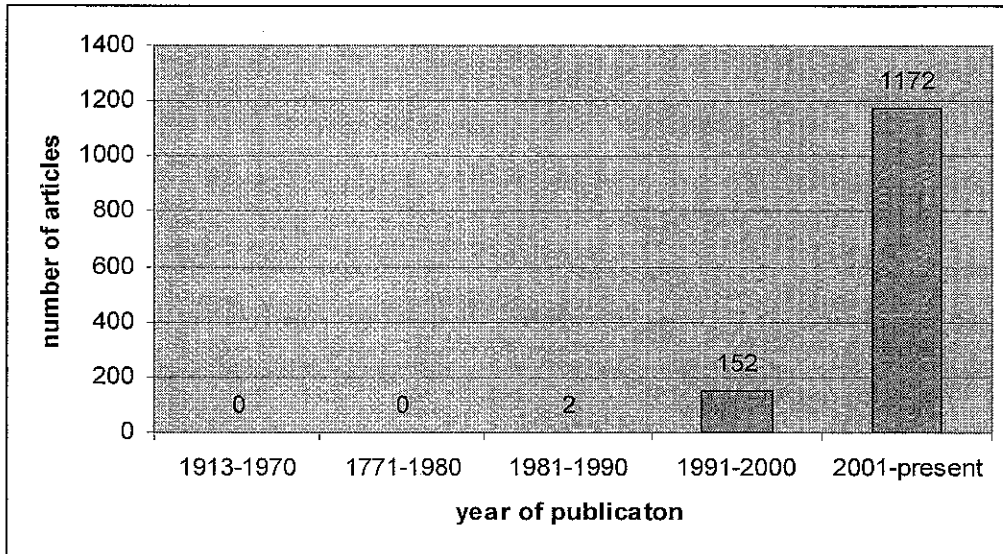


Figure 2.4: Articles found in the IEEE database using keyword image fusion

The first multiresolution based image fusion approach was proposed by [P. J. Burt et al, 1984] using the Laplacian Pyramid as the image code. The algorithm follows a “choose max” selection rule that, at each sample position in the pyramid, the source pyramid coefficient with the maximum value is copied to the composite pyramid. In this implementation let  $A$  and  $B$  represent two input images to be combined; and let  $Y_A$  and  $Y_B$  represent the corresponding Pyramid decompositions for images  $A$  and  $B$  respectively, then, the composite image  $Y_F$  is obtained by:

$$Y_F^k(n, p) = \begin{cases} Y_A^k(n, p) & \text{if } |Y_A^k(n, p)| > |Y_B^k(n, p)| \\ Y_B^k(n, p) & \text{otherwise} \end{cases} \quad (2.10)$$

$Y_F^k(n, p)$  represents the fused detail coefficient at location  $n$  within level  $k$  and band  $p$ .

Toet [Toet A., 1989] proposed another method using a ratio of low pass pyramid and the same “choose max” selection rule which is used by Burt to fuse IR and visible image. These methods generally appear to provide good results. One limitation is in the fusion of patterns that have roughly equal salience but opposite contrast. This is a pathological case since “choose max” results in pattern cancellation, and the selection is unstable.

Burt and Kolczynski [P. J. Burt, and Kolczynski, 1993] proposed another method which introduces the concept of activity measure and match measure. Nowadays this method is taken as a standard for image fusion algorithms. Activity measure is a measure of the saliency or the importance of each pixel while match measure measures the similarity between the two source images. In this work, the activity measure is computed as a local energy measure.

$$a_A^k(n, p) = \sum_{\Delta n \in D_p^k} |Y_A^k(n + \Delta n, p)|^2 \quad (2.11)$$

and the match between  $Y_A$  and  $Y_B$  is computed as a normalized correlation average between the two images as:

$$m_{AB}^k(n, p) = \frac{2 \sum_{\Delta n \in D_p^k} Y_A^k(n + \Delta n, p) Y_B^k(n + \Delta n, p)}{a_A^k(n, p) + a_B^k(n, p)} \quad (2.12)$$

where  $D_p^k$  is the window at level  $k$  and band  $p$  and has size  $1X1$ ,  $3X3$ , or  $5X5$ . The combination process is the weighted average.

$$Y_F^k(n, p) = w_A(d^k(n, p))Y_A^k(n, p) + w_B(d^k(n, p))Y_B^k(n, p) \quad (2.13)$$

where the weights are determined by the decision process for each level  $k$ , band  $p$  and location  $n$  as:

$$w_A(d^k(n, p)) = 1 - w_B(d^k(n, p)) \quad (2.14)$$

with the decision map,  $d$  given by equation 2.15:

$$d^k(n, p) = \begin{cases} 1 & \text{if } m_{AB}^k(n, p) \leq T \text{ and } a_A^K(n, p) > a_B^k(n, p) \\ 0 & \text{if } m_{AB}^k(n, p) \leq T \text{ and } a_A^K(n, p) \leq a_B^k(n, p) \\ \frac{1}{2} + \frac{1}{2} \left( \frac{1 - m_{AB}^k(n, p)}{1 - T} \right) & \text{if } m_{AB}^k(n, p) \geq T \text{ and } a_A^K(n, p) > a_B^k(n, p) \\ \frac{1}{2} - \frac{1}{2} \left( \frac{1 - m_{AB}^k(n, p)}{1 - T} \right) & \text{if } m_{AB}^k(n, p) \geq T \text{ and } a_A^K(n, p) \leq a_B^k(n, p) \end{cases} \quad (2.15)$$

for some match measure (similarity) threshold  $T$ . In case of a poor match, the source coefficient having the largest activity will yield the composite value; otherwise, a weighted sum of the sources coefficients will be used. The authors claim that this approach provides a partial solution to the problem of combining components that have opposite contrast, since such components are combined by selection. Gemma Piella Fenoy [Gemma Piella, 2003] and Li Tao [Li Tao, 2005] used this method their thesis applying it in various wavelet domains. We also use this technique for the fusion process with slight modifications and in contourlet domain.

Ranchin and Wald [Ranchin T. and Wald L., 2000] presented one of the first wavelet-based fusion systems. Their implementation considers the maximum absolute value within a window as the activity measure associated with the sample centered in the window:

$$a_s^k(n, p) = \max \left( \left| Y_s^k(n + \Delta n, p) \right| \right)_{\Delta n \in W_p^k} \quad (2.16)$$

For each position in the transform domain, the maximum selection rule is used to determine which of the inputs is likely to contain the most useful information. This results in a preliminary decision map which indicates, at each position, which source should be used in the combination map. They claimed that they have found better fidelity.

Li and Wang [S. T. Li and Y.N. Wang, 2000] examined the application of discrete multiwavelet transform to multi-sensor image fusion. The composite coefficients are obtained through a sample-based maximum selection rule. The authors showed experimental results where their fusion scheme performs better than those based on comparable scalar wavelet transforms.

Several other fusion methods exist which use the multiresolution using the wavelet transform [R. Yang et al 2006, Yi Yang et al 2007, K. Amolins et al 2007, S.G Nikolov et al 1999, H. Wang 2004, Yin Chen 2007]. However, wavelet bases present some limitations, because they are not well adapted to the detection of highly anisotropic elements, such as alignments in an image, or sheets in a cube. In this thesis we propose a generalized image fusion technique which uses the multiresolution by contourlet transform. The selection method is a modified version of the Burt and Kolczynski [Burt P. J, and Kolczynski, 1993] algorithm. In our approach the approximate and detail contourlet coefficients are treated in different ways as they carry different physical meanings. Our technique is presented in detail in chapter six. For detail investigation of existing image fusion algorithms the reader is referred to the PhD thesis by Yin Chen [Yin Chen, 2007].

The next chapter will look into the adaptation of transformation domain for image processing and mathematical formulation of the new multiresolution technique (contourlet transform).

## CHAPTER THREE: MULTIREOLUTION DECOMPOSITION

### 3.1 Introduction

There exist several ways to represent a signal and the efficiency of a given representation depends on the subsequent processing. Different representations emphasize different aspects of a signal and therefore, one should look for representations that make relevant information easily accessible.

Non linear approximation (NLA) techniques have played a major role in approximating signal features. In the last two decades, one of the NLA techniques which is called multiresolution representation occupy an important role and have proved to be a powerful tool both theoretically and practically. This chapter provides a brief overview of two multiresolution representations, namely the wavelet transform and contourlet transform.

A signal structure depends on the scale at which the signal is being perceived, it should be analyzed at different scales or levels of resolution. This is called multiresolution representation.

Multiresolution is actually a biologically inspired technique [S. Mallat, 1999]. Consider for example how our eyes see an object. Let us consider a forest scene at different scales. As we get closer, we can distinguish individual trees, then branches, and finally the leaves. As we go for smaller and smaller scales, we can see details that we did not see before, but at the same time the view field is reduced. Some theories involving the way mammalian brains process auditory and visual information suggest that human perception can be modeled with multiresolution analysis [S. Mallat, 1999, E. Hernandez and G. Weiss, 1996].

The following are some of the several reasons why multiresolution approach is preferable over other approaches [S. Mallat, 1999, E. Hernandez and G. Weiss, 1996].

- Most signals exhibit relevant features at many different resolutions.
- Human visual system process information in multiresolution fashion.
- Sensors may provide the raw data in different resolution.
- We may need for output in different resolution levels.
- It may also offer computational advantages.

Recent developments in multiresolution representation approaches include wavelet, ridgelet, curvelet and contourlet transforms. Most researchers indicate that directionality is a crucial feature in image representation; so we have chosen the contourlet transform as our major image decomposition method which has proved to be a true 2D transform. In this thesis, we will mainly discuss wavelet and contourlet representations, there are, however, many other MR techniques. They have the same multiresolution paradigm. But apart from that they differ in many respects, both in theory and in practice. In the next subsections we describe the wavelet and contourlet approaches.

### 3.2 Wavelet Transform

Wavelets are mathematical functions that are defined over finite intervals and have an average value of zero. They transform data into different frequency components representing each component with a resolution matched to its scale.

The basic idea of the wavelet transform is to represent an arbitrary signal as a superposition of a set of wavelets or basis functions. This basis functions are obtained from a single prototype wavelet called the *mother wavelet*, by dilation or contraction (scaling) and translation (shift).

Wavelets have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. Many wavelet

applications such as image compression, image segmentation, and image enhancement are developed in recent years [M. Vetterli, 2001].

The wavelets form a family, where the basic form is called the *mother wavelet*. All the daughter wavelets are derived from this wavelet according to the equation 3.1 by time shifting and time scaling. These daughter wavelets are usually called basis functions.

$$\Psi_{s,\tau} = \frac{1}{\sqrt{s}} \Phi\left(\frac{t-\tau}{s}\right) \quad (3.1)$$

The term  $\frac{1}{\sqrt{s}}$  normalizes the energy for different scales, whereas the  $s$  and  $\tau$  define width and translation of the wavelet respectively. The mother wavelet is defined from equation 3.1 when  $s=1$  and  $\tau=0$

$$\Psi_{1,0}(t) = \Phi(t) \quad (3.2)$$

The continuous wavelet transform (CWT) as a function of  $f(t)$  is defined as:

$$F(s, \tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} f(t) \Phi\left(\frac{t-\tau}{s}\right) dt \quad (3.3)$$

The CWT can be interpreted in two different ways [W.J. Phillips, 2003]:

1. The CWT is the *inner product* or cross correlation of the signal  $f(t)$  with the scaled and time shifted wavelet  $\frac{1}{\sqrt{s}} \Phi\left(\frac{t-\tau}{s}\right)$ . This cross correlation is the measure of the similarity between the signal and the shifted and scaled version of the wavelet.

Note: The *inner product* of two signals  $x(t)$  and  $y(t)$  which is defined in the interval  $a < t < b$  is:

$$\text{inner product} = \langle x(t), y(t) \rangle = \int_a^b x(t)y(t)dt \quad (3.4)$$

2. For a fixed scale  $s$ , the CWT is the convolution of the signal with time reversed

$$\text{wavelet } w(t) = \frac{1}{\sqrt{s}} \Phi\left(\frac{-t}{s}\right).$$

To clearly understand the wavelet transform, let us describe the generalized analysis and synthesis theorem as stated in Phillips lecture notes [W.J. Phillips, 2003]. This theorem says:

*“Given an orthogonal basis signals  $\Phi_n(t)$ , we can analyze any signal  $f(t)$  in terms of the basis  $\Phi_n(t)$  and perfectly synthesize the signal back again.”*

If the scales and positions are chosen based on powers of two, the so called dyadic scales and positions, then calculating wavelet coefficients are efficient and accurate. In CWT we can get an orthogonal basis function by choosing the scales to be powers of two and the times to be an integer multiple of the scales. For instance, let  $j$  and  $k$  be integers, then the CWT can be rewritten as:

$$F\left(\frac{1}{2^j}, \frac{k}{2^j}\right) = 2^{\frac{j}{2}} \int_{-\infty}^{\infty} f(t) \Phi(2^j t - k) dt \quad (3.5)$$

where  $s = \frac{1}{2^j}$  and  $\tau = \frac{k}{2^j} = ks$ .

Now we can define a set of daughter wavelets by double indexing as follows:

$$\Phi_{j,k}(t) = 2^{\frac{j}{2}} \Phi(2^j t - k) \quad (3.6)$$



$$F\left(\frac{1}{2^j}, \frac{k}{2^j}\right) = 2^{\frac{j}{2}} \int_{-\infty}^{\infty} f(t) \Phi_{j,k}(t) dt \quad (3.7)$$

Define a variable  $C_{j,k} = F\left(\frac{1}{2^j}, \frac{k}{2^j}\right)$  and they are called the analysis coefficients. There exist large classes of wavelet functions for which the set of daughter wavelets are an orthogonal basis. Morlet, Daubachies, Symlet, Haar, are some of these wavelets. One can also design his own wavelets to suite the requirements of his/her situation. Daubachies wavelets are the most common and we mainly use them in our wavelet analysis.

In the case of an orthogonal wavelet the analysis formula is called the Discrete Wavelet Transform (DWT). This is the one which is employed in the digital image processing applications. We can rewrite the formula as,

$$C_{j,k} = 2^{\frac{j}{2}} \int_{-\infty}^{\infty} f(t) \Phi_{j,k}(t) dt \quad (3.8)$$

The recovery of the signal through the synthesis formula given in equation 3.9 is called the Inverse Discrete Wavelet Transform (IDWT).

$$f(t) = \sum_j \sum_k C_{j,k} \Phi_{j,k} \quad (3.9)$$

### 3.2.1 Multiresolution Analysis and Filter Banks for Wavelet Transform

Mallat and Meyer [S. Mallat, 1989, Y. Meyer, 1992] introduced the concept of multiresolution analysis. The main idea is by using the orthogonal set wavelets we can easily analyze and perfectly recover a given signal. The use of such kind of wavelets give rise to the multiresolution analysis.

Call  $W_j$  to be the set of orthogonal spaces of all functions  $f(t)$  that can be synthesized from the wavelets  $\Phi_{j,k}(t)$ , where  $-\infty < k < \infty$ , i.e.

$$f(t) = \sum_{j=-\infty}^{\infty} x_j(t) \quad (3.10)$$

where  $x_j(t) = \sum_{k=-\infty}^{\infty} C_{j,k} \Phi_{j,k}(t)$

Note that  $x_j(t)$  is in the space  $W_j$ . Again define  $V_j$  to be the set of all orthogonal spaces of all functions  $f(t)$  that can be synthesized from the wavelets  $\Phi_{i,k}(t)$  where  $i < j$  and  $-\infty < k < \infty$ , then we have:

$$f(t) = \sum_{i=-\infty}^{j-1} C_{i,k} \Phi_{i,k}(t) \quad (3.11)$$

The spaces  $V_j$  are nested inside each other,

$$V_j \subset V_{j-1} \quad j \in \mathbb{Z} \quad (3.12)$$

The nested spaces in equation 3.12 can be visualized by Figure 3.1.

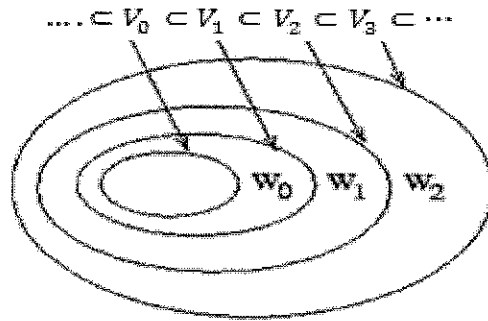


Figure 3.1 Visualization of Nested Subspaces

Every signal in  $V_{j+1}$  is then a sum of a signal in  $V_j$  and  $W_j$ , because:

$$f(t) = \sum_{i=-\infty}^j \sum_k C_{i,k} \Phi_{i,k}(t) = \sum_{i=-\infty}^{j-1} \sum_k C_{i,k} \Phi_{i,k}(t) + \sum_k C_{j,k} \Phi_{j,k}(t) \quad (3.13)$$

Thus we have:

$$V_{j+1} = V_j + W_j \quad (3.14)$$

This shows that the spaces  $W_j$  are the differences of between adjacent spaces of  $V$ :

$$W_j = V_{j+1} - V_j \quad (3.15)$$

To clarify this analysis let us decompose a signal  $f(t)$  in the space  $V_0$ . First the space  $V_0$  can be decomposed as [W.J. Phillips, 2003]:

$$\begin{aligned} V_0 &= V_{-1} + W_{-1} \\ &= V_{-2} + W_{-2} + W_{-1} \\ &= V_{-3} + W_{-3} + W_{-2} + W_{-1} \\ &= V_{-4} + W_{-4} + W_{-3} + W_{-2} + W_{-1} \end{aligned}$$

This leads us to the following decomposition of the signal  $f(t)$

$$\begin{aligned} f(t) &= A_1(t) + D_1(t) \\ &= A_2(t) + D_2(t) + D_1(t) \\ &= A_3(t) + D_3(t) + D_2(t) + D_1(t) \\ &= A_4(t) + D_4(t) + D_3(t) + D_2(t) + D_1(t) \end{aligned}$$

where  $D_i(t)$  in the space  $W_{-i}$  and  $A_i(t)$  in the space  $V_{-i}$  are called detail and approximate coefficients at level  $i$  respectively. For many signals, the low frequency content is the most important part. It is the identity of the signal. The high frequency content, on the other hand, imparts details to the signal. In wavelet analysis, the

approximations and detail are obtained after filtering and sampling processes. The approximations are the high scale, low frequency components of the signal. The details are the low scale, high frequency components.

Calculating wavelet coefficients at every possible scale is a fair amount of work, and it generates an awful lot of data. The space  $V_j$  will be very big as  $j \rightarrow \infty$ , thus it is a common practice to approximate a signal  $f(t)$  by closely choosing  $j = J$  which is large enough.

Using the basis  $\Phi_{J,k}(t)$  (for all values of  $k$ ), the approximation  $A_0$  is given by:

$$A_0(m) = \int_{-\infty}^{\infty} f(t) \Phi_{J,m}(t) dt \quad m \in \mathbb{Z} \quad (3.16)$$

From these coefficients we can approximately reconstruct the original signal,

$$f(t) \approx \sum_m A_0(m) \Phi_{J,m}(t) \quad (3.17)$$

After this approximation in space  $V_J$ , we can further decompose it by using sub spaces  $V_{J-n}$  and  $W_{J-n}$  with their basis  $\Phi_{J-n,k}(t)$ . Each signal can be expressed in two ways using the basis functions in each of the spaces,

$$\begin{aligned} f(t) &= \sum_k A_0(k) \Phi_{J,k}(t) \\ &= \sum_k A_1(k) \Phi_{J-1,k}(t) + \sum_k D_1(k) \Phi_{J-1,k}(t) \end{aligned} \quad (3.18)$$

As the wavelets and spaces at each level are orthogonal, we can compute the coefficients  $A_1(k)$  and  $D_1(k)$  by using inner products as follows:

$$A_1(k) = \langle f(t), \Phi_{j-1,k}(t) \rangle \quad (3.19)$$

Substituting the approximation from equation (3.17) in place of the signal  $f(t)$  in equation (3.19),

$$\begin{aligned} A_1(k) &= \left\langle \sum_n A_0(n) \Phi_{j,n}(t), \Phi_{j-1,k}(t) \right\rangle \quad n \in Z \\ &= \sum_n A_0 \langle \Phi_{j,n}(t), \Phi_{j-1,k}(t) \rangle \end{aligned} \quad (3.20)$$

Solving the inner product, and using equation 3.6:

$$\langle \Phi_{j,n}(t), \Phi_{j-1,k}(t) \rangle = \int_{-\infty}^{\infty} \sqrt{2^j} \Phi(2^j t - n) \sqrt{2^{j-1}} \Phi(2^{j-1} t - k) dt \quad (3.21)$$

Let  $s = 2^{j-1} t - k$  then equation 3.21 reduces to:

$$\langle \Phi_{j,n}(t), \Phi_{j-1,k}(t) \rangle = \int_{-\infty}^{\infty} \sqrt{2} \Phi(2s + 2k - n) \Phi(s) ds \quad (3.22)$$

Using the two scale equation theorem as defined in Phillips lecture notes [W.J. Phillips, 2003] says:

*“Satisfying certain conditions, there exist discrete time filter coefficients  $h(n)$  which lead the basis function  $\Phi(t)$  to be expressed as  $\Phi(t) = \sum_n h(n) \sqrt{2} \Phi(2t - n)$ .”*

Applying this theorem we can further simplify equation 3.22 to:

$$\begin{aligned}
\langle \Phi_{j,n}(t), \Phi_{j-1,k}(t) \rangle &= \int_{-\infty}^{\infty} \sqrt{2} \Phi(2s + 2k - n) \sum_m h(m) \sqrt{2} \Phi(2s - m) ds \\
&= \sum_m h(m) \int_{-\infty}^{\infty} \sqrt{2} \Phi(2s + 2k - n) \Phi(2s - m) ds
\end{aligned} \tag{3.23}$$

Due to the orthogonality of the bases the integral is zero unless and otherwise  $m = 2k - n$ , thus equation 3.23 reduces to:

$$\langle \Phi_{j,n}(t), \Phi_{j-1,k}(t) \rangle = h(n - 2k) \tag{3.24}$$

Therefore, the approximate coefficients are given by:

$$A_1(k) = \sum_n h(n - 2k) A_0(n) \tag{3.25}$$

With a similar argument the detail coefficients are given by using another filter  $g(n)$ :

$$D_1(k) = \sum_n g(n - 2k) A_0(n) \tag{3.26}$$

These two equations resemble the convolution followed by downsampling and give a rise to the filter bank representation of the multiresolution analysis. Downsampling a discrete signal  $x(n)$  consists of omitting every other value as shown in Figure 3.2,

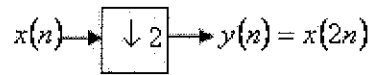


Figure 3.2 Downsampling

Thus if we follow a filter by a downsampler, we will get the approximate coefficients at the next level as shown in Figure 3.3.



Figure 3.3 Filtering and downsampling for approximate coefficients

The detail coefficients can also be generated by using filter bank arrangement ( Figure 3.4) ,



Figure 3.4 Filtering and downsampling for detail coefficients

Combining the above two filters gives us a one stage analysis filter bank shown in Figure 3.5.

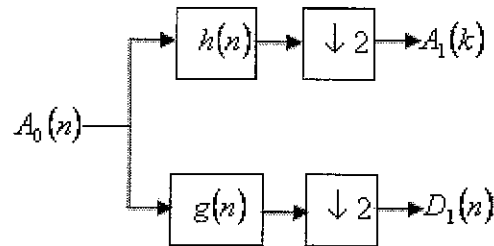


Figure 3.5 One stage analysis filter bank

The coefficients  $A_m(k)$  and  $D_m(k)$  for  $m = 1, 2, 3, \dots$  can be computed by iterating or cascading the single stage filter bank to obtain multistage filter bank as shown in Figure 3.6.

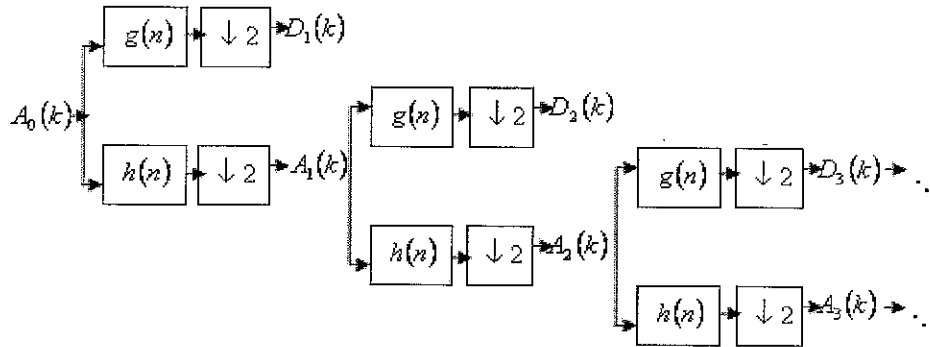


Figure 3.6 Multiresolution Filter Bank Decomposition

As we said earlier, the decomposition of a signal into an approximate and detail coefficients is reversible. This can also be done by using the upsampling and filtering process. The upsampling of a discrete signal consists of inserting zeros between the values as shown in Figure 3.7.

$$x(n) \rightarrow \boxed{\uparrow 2} \rightarrow y(n) = \begin{cases} x\left(\frac{n}{2}\right) & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd} \end{cases}$$

Figure 3.7 Upsampling

The one stage synthesis filter bank is given by Figure 3.8.

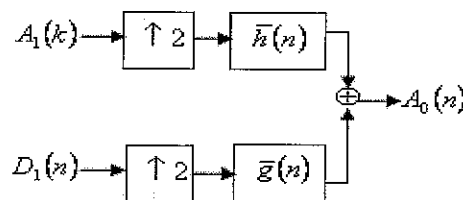


Figure 3.8 One stage Upsampling

The outputs of the multistage analysis filter banks can be fed into a multistage synthesis filter banks to reproduce the original signal. Figure 3.9 shows a 3 stage synthesis filter bank:



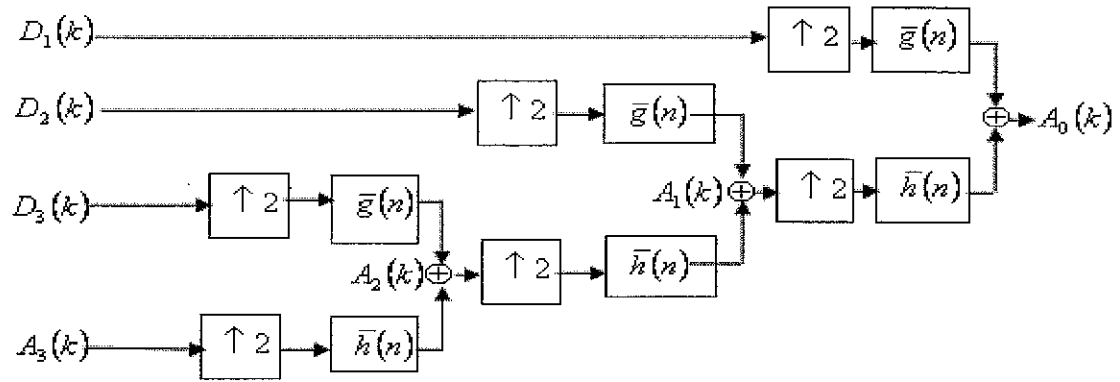


Figure 3.9 A three stage synthesis filter bank

Most wavelet transforms are designed in one dimensional case. By successive application of such one dimensional transforms on the rows and the columns (or vice versa) of an image, one obtains the so called two dimensional wavelet transform. Thus 2D wavelets are found by tensor products of 1D wavelets. This construction is illustrated in Figure 3.10 below. At each level  $k$ , the input  $x(k)$  is decomposed into a coarse approximation  $x(k+1)$  and three detail signals  $y(k+1) = \{y_1(k+1), y_2(k+1), y_3(k+1)\}$  corresponding to horizontal, vertical and diagonal directions.

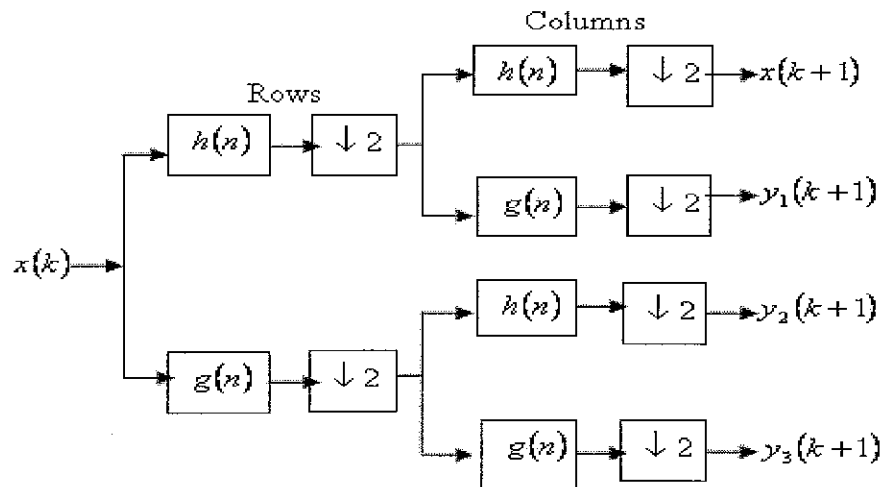


Figure 3.10 A one level 2D decomposition wavelet filter bank

### 3.3 Contourlet Transform

Over the last two decades, wavelets have had a growing impact on signal processing, mainly due to their good *non linear analysis* performance for piecewise smooth functions in one dimension [S. Mallat 1999, M. Vetterli 2001, R. A. Devore et al 1992]. Unfortunately, this is not always true in two dimensional signals. Wavelets are good at catching point discontinuities (zero dimension), but two dimensional piecewise smooth functions which resemble images have one dimensional discontinuity. Intuitively, wavelets in 2D obtained by tensor products of one dimensional wavelets will be good at isolating the discontinuities at edge points, but will not see the smoothness along the contours. This indicates that a more powerful representation is needed in higher dimension.

Recently, Cades and Donho [ E.J. Cades and D.L. Donho 1999, E.J. Cades and D.L. Donho 2000] established a new system of representation named, *curvelet transform*, that has shown to achieve optimal approximation for 2D piecewise smooth functions. It was defined in the continuous domain and it is proved mathematically that the continuous form of the curvelet transform is rotation invariant and the expansion to the curvelet transform can produce perfect reconstruction. But the curvelet transform poses problems when one translates it into the discrete world. Since it is defined in the polar coordinate the implementation of curvelet transform for discrete images on rectangular coordinates is very challenging.

With this insight, M. N. Do and M. Vetterli [M. N. Do and M. Vetterli ,2005] developed the *contourlet transform* in discrete domain which is defined on rectangular coordinates instead of polar coordinates and it is more computers friendly. It offers a flexible multiresolution and directional decomposition for images, since it allows different number of directions at each scale.

### 3.3.1 Theory of Contourlets

The essential goal of contourlet construction is an improvement over the wavelets as wavelets lack directionality and they are blind in capturing the image important features as lines, edges, curves and contours. To comprehend how one can improve the 2D wavelet transform using the contourlet transform for representing images with smooth contours from M. N. Do and M. Vetterli work, consider the following scenario as illustrated in Figure 3.11.

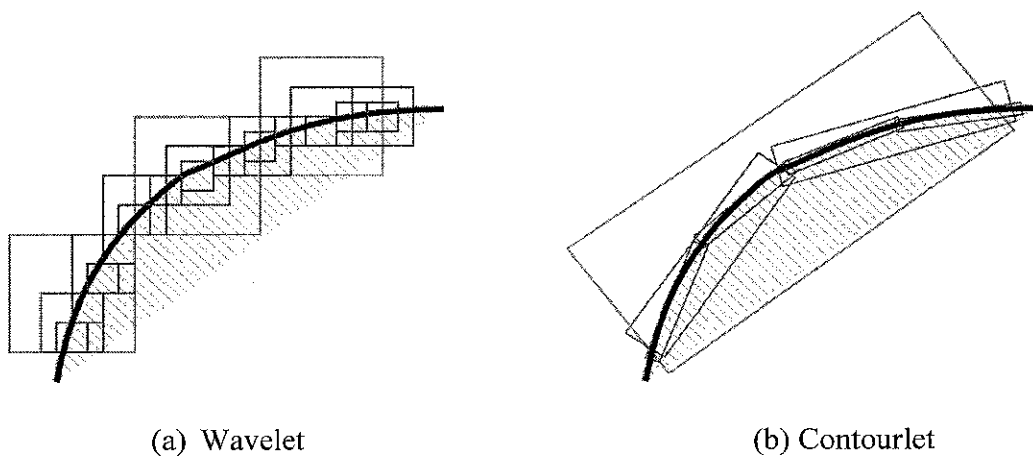


Figure 3.11 Wavelet vs. Contourlet

Imagine representing a curve, one with a wavelet style and the other with a contourlet style, both cases wish to reproduce it as a natural scene. We apply a refinement technique to increase resolution from course to fine. Here, efficiency is measured by how quickly one can faithfully reproduce the scene.

By definition the wavelet style is limited to use square shaped bases along the contour, using different sizes corresponding to the multiresolution structure of wavelets. As the resolution becomes finer, we can clearly see the limitation of the wavelet style which needs to use many fine dots to capture the contour. The contourlet style on the other

hand, exploits effectively the smoothness of the contour by making bases with different elongated shapes and in a variety of directions following the contour.

Comparing the wavelet scheme with the contourlet scheme in Figure 3.11, we see that the improvement of contourlets can be interpreted as grouping of nearby wavelet coefficients, since their location are locally correlated due to the smoothness of the discontinuity curve [M. N. Do and M. Vetterli, 2005]. Therefore, we can obtain a sparse image expansion by first applying a multi-scale transform and then applying a local directional transform to gather the nearby basis functions at the same scale into linear structures.

With this insight, a double filter bank approach for obtaining sparse expansion for typical images with smooth contours is developed by M. N. Do and M. Vetterli. The Laplacian pyramid (LP) [P. J. Burt and E. H. Adelson, 1983] is first used to capture point discontinuities, and then followed by a directional filter bank (DFB) [R.H. Bamberger and M.J. T. Smith, 1992] to link point discontinuities into linear structures. The over all result is an image expansion using elementary images like contour segments, and thus it is called Pyramidal directional filter bank decomposition (PDFB) or the contourlet transform.

### 3.3.2 Laplacian Pyramid

The Laplacian pyramid (LP), first introduced by Burt and Adelson [P.J. Burt and E.H. Adelson, 1983] is the decomposition of the original image such that each level corresponds to a different band of image frequencies. Figure 3.12 depicts one level decomposition process, where  $H$  and  $G$  are called analysis and synthesis filters respectively and  $M$  is the sampling matrix.

The basic idea behind LP is, first determine coarse approximation  $c$ , by filtering and down sampling, then find the prediction  $p$ , by upsampling and filtering. Finally determine

the difference  $d$ , between the prediction  $p$  and original  $x$ . The process can be iterated on the coarse (down sampled low pass) signal.

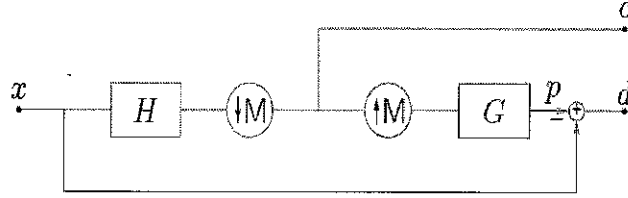


Figure 3.12 One level Laplacian Pyramid decomposition

The coarse approximation  $c$  signal is thus given by,

$$c[n] = \sum_{k \in \mathbb{Z}} x[k] h[Mn - k] \quad (3.27)$$

where  $z$  is an integer.

The predictions  $p$ , signal is given as:

$$p[n] = \sum_{k \in \mathbb{Z}} c[k] g[n - Mk] \quad (3.28)$$

Finally the difference signal,  $d$ , can be written as a matrix subtraction as follows,

$$d[n] = x[n] - p[n] \quad (3.29)$$

Figure 3.13 shows one level reconstruction scheme for the Laplacian Pyramid. The constructed signal  $\hat{x}$  is found by adding back the difference,  $d$  to the prediction from the coarse signal  $c$ .

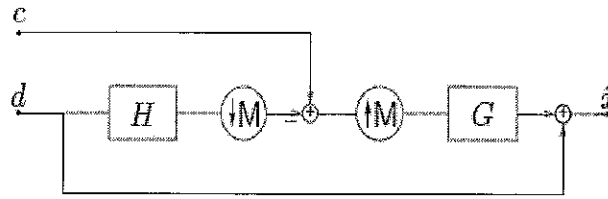


Figure 3.13 One level Laplacian Pyramid Reconstruction

Over sampling is an implicit drawback of the LP [M. N. Do and M. Vetterli, 2005]. Therefore in compression applications LP is completely replaced by wavelet transform. However, in contrast to the critically sampled wavelet scheme, the LP has the distinguishing feature that each pyramid level generates only one band pass image. This makes it very easy to apply in many multiresolution applications. And the band pass image does not have scrambled frequencies. This *frequency scrambling* happens in the wavelet filter bank when a high pass channel, after down sampling, is folded back into the low frequency band and its spectrum is reflected. In the LP, this effect is avoided by down sampling the low pass channel only. Thus the band pass output of the LP can be further decomposed in directional *filter banks*.

### 3.3.3 Directional Filter Bank Decomposition

Bamberger and Smith pioneered a 2D directional filter bank [R.H. Bamberger and M.J. T. Smith, 1992] that can be maximally decimated while achieving perfect reconstruction. They implemented in a tree structure composed of two band systems. The signal is modulated by  $\pi$  to maximally decimate the output. Consider an 8-band wedge-shaped frequency decomposition of a signal as shown in Figure 3.14. Wedges 0, 1, 2, and 3 are mainly horizontal and wedges 4, 5, 6 and 7 are mainly vertical.

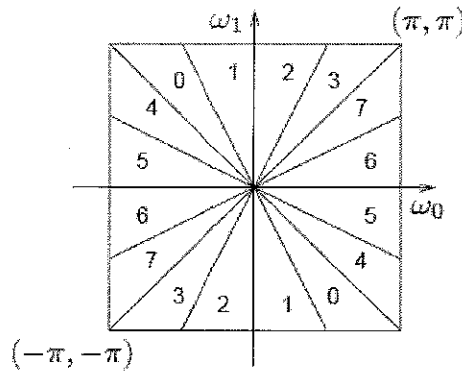


Figure 3.14 Directional filter bank frequency partitioning where  $l=3$  and there are  $2^l = 8$  wedge shaped frequency bands.

The original construction of the DFB in [R.H. Bamberg and M.J. T. Smith, 1992] involves modulating the input signal and using diamond shaped filters. Furthermore, to obtain the desired frequency partition, a complicated tree expanding rule has to be followed.

A simpler rule for expanding the decomposition tree and which avoids the modulation process is proposed in [M.N. Do 2001, M.N. Do and Martin Vetterli 2001, M.N. Do and Martin Vetterli 2002]. This method is constructed from two building blocks. The first block is quincunx filter with fan filters which divides the input spectrum into vertical and horizontal directions (see Figure 3.15). The second block is considered as a shearing operator, which actually does only a reordering of the input image, see Figure 3.16 for shearing operator example. Intuitively, the wedge-shaped frequency partition (Figure 3.14) of the DFB is realized by an appropriate combination of directional frequency splitting by the fan Quincunx filter bank (QFB) [Yi Chen et al, 2006] which generated horizontal and vertical decomposition, and the rotation operations done by resampling.

The *quincunx filter* bank, designated by  $Q$ , serves as the core of the directional filter

bank. The two possible representations of the quincunx filter are:  $Q_0 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

and  $Q_1 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ .

$Q_0$  rotates the input image by  $-45^\circ$  and  $Q_1$  rotates it by  $45^\circ$ . Both can be used alternatively.

*Unimodular integer matrices* can be used as a special sampling operation. And they are

usually denoted by  $R$ . We have 4 of them namely:  $R_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ ,  $R_1 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ ,

$R_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$  and  $R_3 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ .

Note that,

$$R_0 R_1 = R_2 R_3 = I \quad (3.30)$$

where  $I$  is an identity matrix. They provide rotation operation to the input image.

Upsampling by  $R_0$  is the same as down sampling by  $R_1$ . Sampling by a unimodular integer matrix does not change the data rate but only rearranges the input samples. We can rewrite the quincunx filter by using the *smith form*.

$$\begin{aligned} Q_0 &= R_1 D_0 R_2 = R_2 D_1 R_1 \\ Q_1 &= R_0 D_0 R_3 = R_3 D_1 R_0 \end{aligned} \quad (3.31)$$

where  $D_0 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$  and  $D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$  diagonal matrices which correspond to dyadic.



The realization of QFB can be given by the diagram in Figure 3.15.

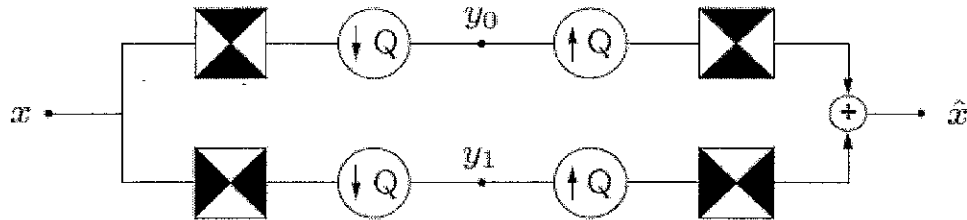


Figure 3.15. Two dimensional spectrum partition using quincunx filter banks with fan filters.

An example of the shearing operation where a  $-45^\circ$  direction edge becomes a vertical edge is displayed by using the cameraman picture as shown in Figure 3.16.



Figure 3.16 Application of a rotation operator where the sampling is done by  $R_0$

### 3.3.4 Multiscale and Directional Decomposition (Contourlets)

Performing the Laplacian decomposition followed by the directional filter bank decomposition produces the contourlet transform framework. The framework is shown in Figure 3.17. The Laplacian Pyramid allows subband decomposition to be applied on an input digital image. Those band pass output images can be fed into a DFB so that directional information can be captured efficiently.

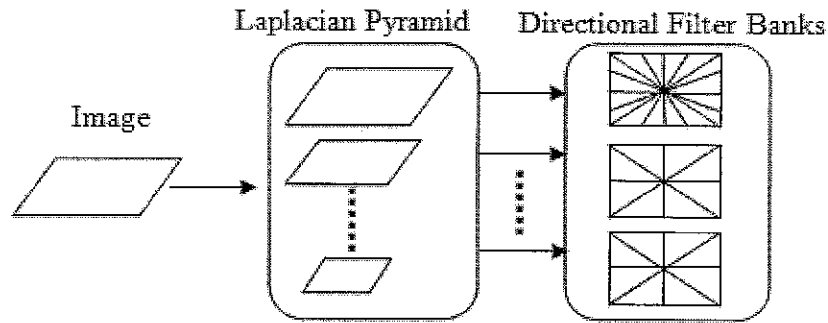


Figure 3.17 Contourlet transform framework

The scheme can be iterated repeatedly on the coarse image as shown in Figure 3.18. The end result is a double iterated filter bank structure; named pyramidal directional filter bank (PDFB) scheme or contourlet transform which decomposes images into directional subbands at multiple scales. The scheme is flexible since it allows for a different number of directions at each scale.

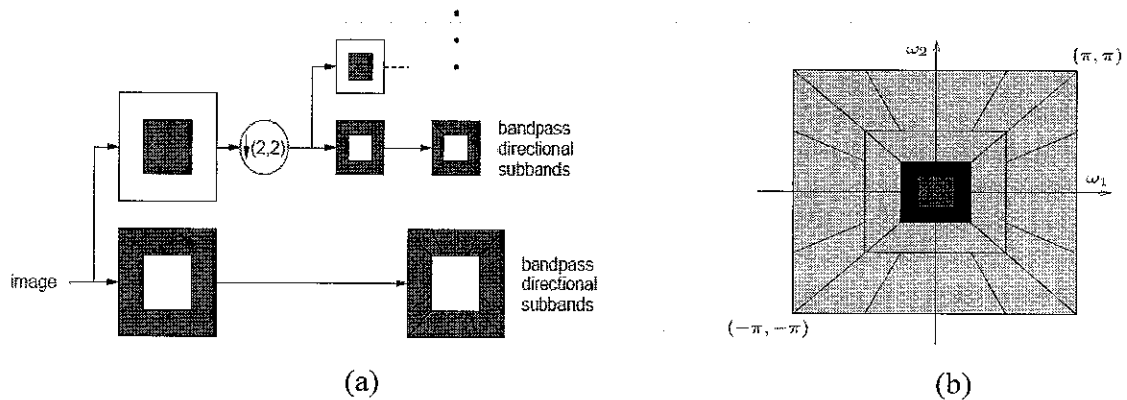


Figure 3.18 a) Contourlet filter bank b) Resulting frequency division

Several researchers are nowadays changing their attention to the contourlet transform and lot of work has been done recently [Duncan Dun and Yin Po 2003, Arthur L. et al 2006, M. N. Do 2003].

### 3.3.5 Multiresolution Analysis and Filter Banks for Contourlet Transform

Bamberger and Smith [R.H. Bamberger and M.J. T. Smith, 1992] said DFB is efficiently implemented via  $l$ -level tree-structured decompositions that lead to  $2^l$  subbands with wedge-shaped frequency partitions. Do and Vetterli [M. N. Do and M. Vetterli, 2005] proposed a new method for DFB by using multi-rate identities. They said  $l$ -level tree structured DFB can be transformed into parallel structures of  $2^l$  channels with equivalent filters and sampling matrices.

Denote these equivalent synthesis filters as  $G_k^{(l)}, 0 \leq k < 2^l$ , which corresponds to the subbands as indexed in Figure 3.14.

The oversampling matrices have diagonal form as:

$$S_k^{(l)} = \begin{cases} \begin{bmatrix} 2^{l-1} & 0 \\ 0 & 2 \end{bmatrix} & \text{for } 0 \leq k < 2^{l-1} \quad \text{near horizontal subbands} \\ \begin{bmatrix} 2 & 0 \\ 0 & 2^{l-1} \end{bmatrix} & \text{for } 2^{l-1} \leq k < 2^l \quad \text{near vertical subbands} \end{cases} \quad (3.32)$$

By translating the impulse response of the synthesis directional filters  $G_k^{(l)}$  over the sampling  $S_k^{(l)}$  we obtain a family of basis for discrete signals as:

$$\{g_k^{(l)}[n - S_k^{(l)}m]\}_{0 \leq k < 2^l, m \in \mathbb{Z}^2} \quad (3.33)$$

As it was said earlier the contourlet elements are multiple directions and the combination of multiscales. To achieve multiscale, suppose the LP in the PDFB uses orthogonal filter and downsampling by two is taken in each direction. Under certain conditions, the low

pass filter  $G$  in the LP uniquely defines an orthogonal scaling function  $\phi(t) \in L^2(\mathbb{R}^2)$  via the two scale equation theorem as stated in [W.J. Phillips, 2003]:

$$\phi(t) = 2 \sum_{n \in \mathbb{Z}^2} g[n] \phi(2t - n) \quad (3.34)$$

$$\phi_{j,n} = 2^{-j} \phi\left(\frac{t - 2^j n}{2^j}\right), j \in \mathbb{Z}, n \in \mathbb{Z}^2 \quad (3.35)$$

The family  $\{\phi_{j,n}\}_{n \in \mathbb{Z}^2}$  is an orthonormal basis of  $V_j$  for all  $j \in \mathbb{Z}$ , where  $V_j$  is a subspace defined on uniform grid intervals  $2^j \times 2^j$ , which characterizes the image approximation at the resolution  $2^{-j}$ .

Let  $W_j$  be the orthogonal complement of  $V_j$  in  $V_{j-1}$  as shown in Figure 3.19.

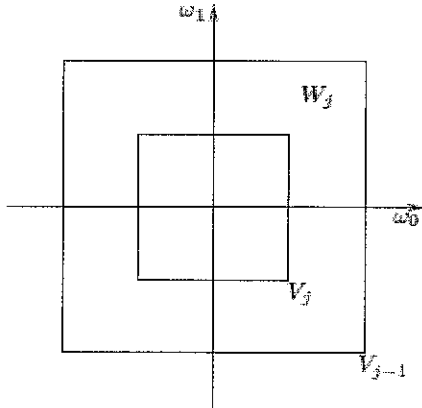
$$V_{j-1} = V_j \oplus W_j \quad (3.36)$$


Figure 3.19 Multiscale nested subspaces generated by the LP

The LP is now an oversampled filter bank where each polyphase component of the difference signal comes from a separate filter bank channel like the coarse signal.

Let  $F_i(z), 0 \leq i < 3$  be the synthesis filter for these polypahse components. As in the wavelet filter bank, we can associate with each of these filters a continuous function  $\varphi^{(i)}(t)$ .

$$\varphi^{(i)}(t) = 2 \sum f_i[n] \phi(2t - n) \quad (3.37)$$

$\mathcal{W}_{j+1}$  is a shift invariant subspace , thus we can have a new variable  $\mu$ :

$$\mu_{j,2n+k_i}(t) = \varphi_{j+1,n}^{(i)} = \sum_{m \in \mathbb{Z}^2} f_i[m] \phi_{j,n+m}(t) \quad (3.38)$$

With this notation, the family  $\{\mu_{j,n}\}_{n \in \mathbb{Z}^2}$  is a tight frame of  $\mathcal{W}_{j+1}$  and it resembles a uniform grid on  $R^2$  of intervals  $2^j \times 2^j$ .

For multiple direction case, suppose that DFB's in the PDFB use orthogonal filters. When DFB is applied to the multiresolution subspace  $V_j$ , on bases  $\{g_k^{(l)}[n - S_k^{(l)}m]\}_{0 \leq k < 2^l, m \in \mathbb{Z}^2}$ , Define:

$$\theta_{j,k,n}^{(l)}(t) = \sum_{m \in \mathbb{Z}^2} g_k^{(l)}[m - S_k^{(l)}n] \phi_{j,m}(t) \quad (3.39)$$

The family  $\{\theta_{j,k,n}^{(l)}\}_{n \in \mathbb{Z}^2}$  is orthonormal bases of a directional subspace  $V_{j,k}^{(l)}$  for each  $k = 0, \dots, 2^l - 1$ . These subspaces are orthogonal as shown in Figure 3.20 with:

$$\begin{aligned} V_{j,k}^{(l)} &= V_{j,2k}^{(l+1)} \oplus V_{j,2k+1}^{(l+1)} \text{ and} \\ V_j &= \bigoplus_{k=0}^{2^l-1} V_{j,k}^{(l)} \end{aligned} \quad (3.40)$$

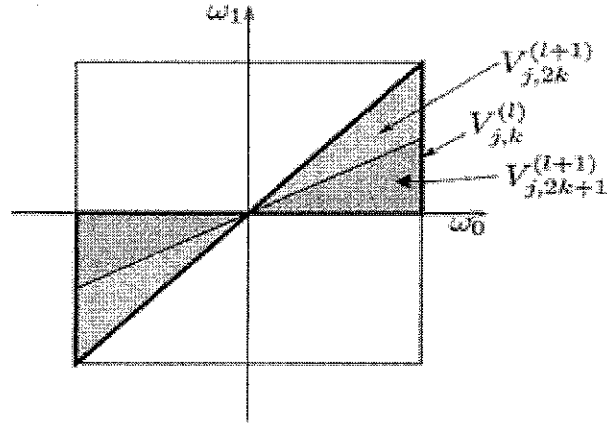


Figure 3.20 Multiresolution nested subspaces generated by the DFB

Applying the directional decomposition by the family  $\{g_k^{(l)}[n - S_k^{(l)}m]\}_{0 \leq k < 2^l, m \in \mathbb{Z}^2}$ , on the detail subspace  $W_{j+1}$  as done by the PDFB, we have a variable  $\rho$ :

$$\rho_{j,k,n}^{(l)}(t) = \sum_{m \in \mathbb{Z}^2} g_k^{(l)}[m - S_k^{(l)}n] \mu_{j,m}(t) \quad (3.41)$$

The family  $\{\rho_{j,k,n}^{(l)}\}_{n \in \mathbb{Z}^2}$  is a tight frame of subspaces  $W_{j+1,k}^{(l)}$  with the frame bound equal to 1 for each  $k = 0, \dots, 2^l - 1$ . These subspaces are orthogonal as shown in Figure 3.21 with:

$$\begin{aligned} W_{j,k}^{(l)} &= W_{j,2k}^{(l+1)} \oplus W_{j,2k+1}^{(l+1)} \quad \text{and} \\ W_j &= \bigoplus_{k=0}^{2^l-1} W_{j,k}^{(l)} \end{aligned} \quad (3.42)$$

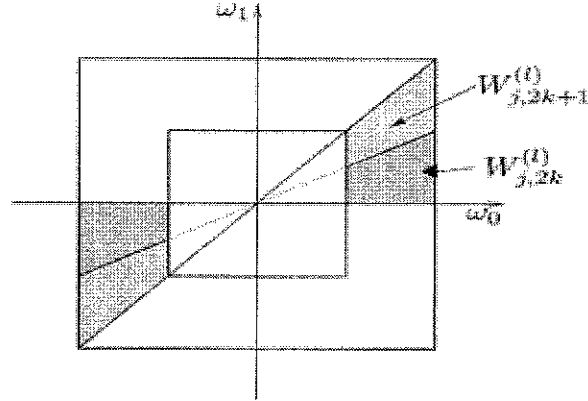


Figure 3.21 Multidirectional nested subspaces generated by the PDFB

Let us denote

$$\rho_{j,k}^{(l)}(t) = \sum_{m \in \mathbb{Z}^2} g_k^{(l)}[m] \mu_{j,m}(t) \quad (3.43)$$

For  $l > 2$

$$\rho_{j,k,n}^{(l)}(t) = \rho_{j,k}^{(l)}(t - 2^j S_k^{(l)} n) \quad (3.44)$$

By substituting equation 3.38 into equation 3.43 we can write the prototype function  $\rho_{j,k}^{(l)}(t)$  directly as a linear combination of the scaling function  $\phi_{j,m}(t)$  and rearranging the summation:

$$\begin{aligned} \rho_{j,k}^{(l)}(t) &= \sum_{i=0}^{l-3} g_k^{(i)}[2n + k_i] \left( \sum_{m \in \mathbb{Z}^2} f_i[m] \phi_{j,n+m}(t) \right) \\ &= \sum_{m \in \mathbb{Z}^2} \left( \sum_{i=0}^{l-3} \sum_{n \in \mathbb{Z}^2} g_k^{(i)}[2n + k_i] f_i[m - n] \right) \phi_{j,m}(t) \end{aligned} \quad (3.45)$$

Let  $c_k^{(l)}[m] = \left( \sum_{i=0}^{l-3} \sum_{n \in \mathbb{Z}^2} g_k^{(i)}[2n + k_i] f_i[m - n] \right)$ , which resembles a summation of convolutions between  $g_k^{(i)}[m]$  and  $f_i[m]$ , thus it is a highpass directional filter. The equation reveals the

contourlet like behavior of the prototype function  $\rho_{j,k}^{(l)}(t)$  where it is seen as grouping of edge detection elements at a scale  $j$  and along a direction  $k$ .

Suppose that  $x[n] = \langle f, \Phi_{L,n} \rangle$  for all  $f \in L^2(R^2)$ , then the discrete contourlet transform of  $x$  is given by:

$$x \xrightarrow{\text{contourlet transform}} (a_J, d_{j,k}^{(l)})_{j=1,\dots,J; k=0,\dots,2^j \times j-1} \quad (3.46)$$

where  $a_J$  and  $d_{j,k}^{(l)}$  are approximate and detail coefficients respectively and are given by:

$$a_J[n] = \langle f, \Phi_{L+J,n} \rangle \quad (3.47)$$

$$d_{j,k}^{(l)}[n] = \langle f, \rho_{j,k}^{(l)} \rangle \quad (3.48)$$

The next chapter looks into the algorithm for single image enhancement approach based on contourlet transform.



## CHAPTER FOUR: SINGLE IMAGE ENHANCEMENT

### 4.1 Introduction

Image enhancement is a process of improving the interpretability or perceptibility of information in images for human viewers or providing better input for other automated image processing techniques. Today there is almost no technical endeavor that is not impacted by digital image enhancement. However, most of the image enhancement algorithms developed and implemented are for only gray scale images despite their rare usages. In this chapter we concentrate on color image enhancement methods. Based on the raw image data representation, image enhancement techniques can be divided into two broad categories:

- Spatial domain methods , which operate directly on pixels ,and
- Transform domain methods, which operate on the transformed image coefficients.

Depending on the number of input images required for enhancement, we can also classify image enhancement methods as single image approaches and composite image approaches. A single image approach takes a single input and enhances that image based on its contents only. This approach is highly dependent on the quality of the image sensor (camera). A composite image approach on the other hand can take several input images to produce a single complete and enhanced image. This technique is usually called image fusion. In this chapter we focus on a single image enhancement approach using transform domain methods.

Unfortunately, there is no general theory for determining what good image enhancement is when it comes to human perception. If it looks good, it is good. However, when image enhancement techniques are used as preprocessing tools for other image processing

techniques, then the quantitative measures can determine which techniques are most appropriate.

In this chapter we propose a new image enhancement algorithm based on contourlet transform. The rich anisotropic and directionality properties of the contourlet transform are tested in digital color image enhancement techniques. The contourlet technique will be compared with two spatial domain methods, histogram equalization (HE) and contrast limited histogram equalization (CLAHE) and the wavelet transform technique.

## **4.2 Multiresolution Based Image Enhancement**

Image enhancement is a necessary step in many image processing applications. The performance of any image processing algorithm is highly dependent on the image representation technique. Images can be processed as a raw data in their spatial domain representation or we can also use non-linear approximation (NLA) techniques to represent the data prior to enhancement. Multiresolution is one of the well known NLA techniques used in image processing applications. It has proved to be a more versatile and efficient method than the currently used spatial methods [Ronald A. Devore, 1998].

Image enhancement is one of the areas where multiresolution is frequently used. It is used to reduce the noise level of a degraded image while sharpening details of the image itself. The basic idea of multiresolution based image enhancement is as follows. Main signal components and noise possess distinct properties in the transform domain. Noises are not parts of the digital images thus they will not generate many coefficients in the transform domain. Especially when we are using the contourlet transform, the effects of the noise are minimal. This is because chance of noises to produce contourlet coefficients is minimal as noises are not parts of the structural information of the image.

The detail coefficients in each level of the multiresolution representation are passed through a specifically designed thresholding (transformation) function where very small

values, which are likely to contain the noise, are set to zero, while medium values are elevated and larger values which include the main signal features are retained. The final enhanced image is obtained by performing inverse transformation process on the modified coefficients.

Low visibility is generally present in images as dark shadows, over bright regions and blurred details. All these are related to the luminance and contrast properties of images; therefore, it is a logical way to develop an image enhancement algorithm based on the processing of the luminance and contrast of an image. Thus the enhancing function should be designed to perform well with low visible luminance images.

For enhancing poorly lit color images, it is therefore required to transform the image into a color space which effectively decomposes the image into luminance and chrominance parts. There exist several color spaces in the literature. Each of them has its own pros and cons. The HSI (Hue, Saturation and Intensity) color model is found to best suited for this purpose (refer to Section 4.5). This is because it can successfully separate the chrominance and luminance component of the image signal.

Therefore, to enhance a color image, the image is first decomposed in to HSI components. The contourlet transform decomposes the intensity,  $I$ , component into sub images based on the number of levels of the specified decomposition. For each sub image, the maximum magnitude coefficient value is computed. Then each sub image is iteratively thresholded/ transformed using the transformation function. Treating each sub image independently has a profound importance. It increases the local contrast of each of the subbands and when the subbands are combined to form the original image the over all or global contrast will be enhanced. The enhancement function is described below.

Let  $I$  represent the subband image in the contourlet or wavelet transformed image; define a value  $x$  as:

$$x = \frac{aI}{\beta M} \quad (4.1)$$

where  $\beta$  is a threshold in which signal features greater than  $\beta$  are retained and usually consists of the above quarter of the image.  $M$  is maximum magnitude of the transformed subband image coefficient.

The enhancement function is then defined as:

$$f(x) = \frac{e^{2bx} - ce^{-x^2} + ce^{2bx-x^2} - 1}{e^{2bx} + 1} \quad (4.2)$$

where  $b$  and  $c$  are parameters which control the gain of the function. The range of  $a$ ,  $b$  and  $c$  are determined empirically based on several enhancement experiments and authors' judgment. To attain optimal enhancement without color distortion and halo effects, the ranges of  $a$  is from 1 to 2.5, for  $b$  is from 1 to 3 and  $c$  is from 1 to 5. Coefficient values greater than  $\beta M$  are enhanced linearly with a gain  $a$ .

The transformation function in equation 4.2 is derived from the combination of the inverse logistic function of sigmoid curve and the hyperbolic tangent ( $\tanh$ ) function. The derivation of this function is given in appendix A.

Then the enhanced image is found by multiplying each coefficient in the subbands by the enhancement function  $f(x)$  as shown below:

$$I_{enhanced} = I \times \text{sign}(I) \times f(x) \quad (4.3)$$

Comparing this function with other multiresolution enhancement methods which use transfer functions,  $y(x)$ , which are defined by using piecewise function combined to become continuous; our method is better in many respects. The methods of Velde [Koen

Vande Velde, 1999] in equation 2.4 using the wavelet transform and the Starck et al [Jean-Luc Starck, 2003] in Equation 2.9 using the curvelet transform are displayed in Figure 4.1(a) and Figure 4.1(b) respectively (For detail explanation of these two systems please refer to Chapter Two, Section 2.1). Using piecewise functions may change the position of the edges and the image content might be degraded.

Therefore, it is a better choice to develop a transformation function that satisfies all the thresholding without using sum of smaller functions. Comparing the shapes of the Velde's method in Figure 4.1(a) and Starck's method in Figure 4.1(b) with our enhancement function displayed in Figure 4.2, our function is smoother and satisfies all the enhancement conditions and has a profound effect in reducing the computational complexity by avoiding several comparison statements in the function implementation.

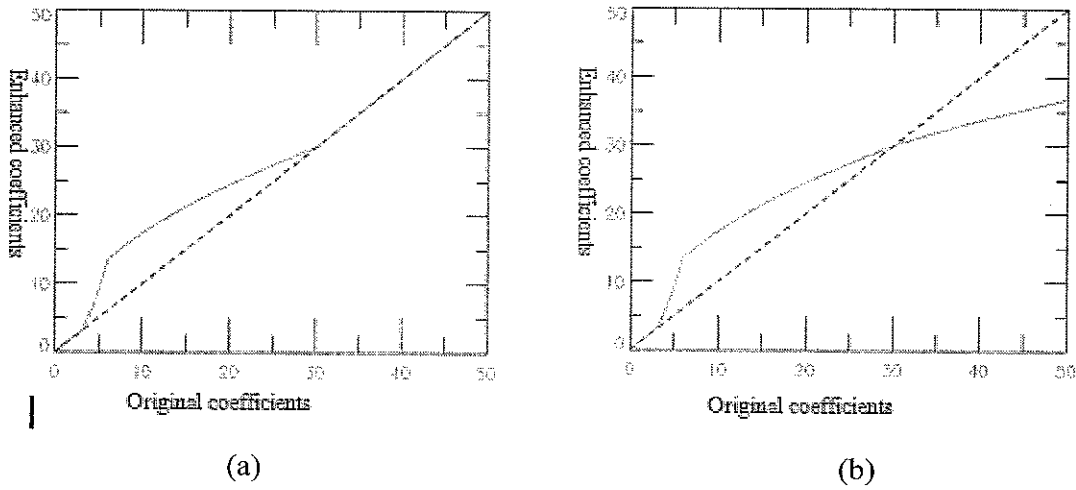


Figure 4.1 Enhanced coefficients versus original coefficients. (a) Velde's method using equation 4.4 with parameters  $m = 30$ ,  $c = 3$ ,  $s = 0$ , and  $p = 0.5$ .

(b) Starck's method using equation 4.5 with parameters  $m = 30$ ,  $c = 3$ ,  $s = 0.6$ , and  $p = 0.5$ .

The transformation function in equation 4.2 satisfies the following fundamental conditions to perform the enhancement without introducing artificial artifacts.

- The function is infinitely differentiable. This is a very important property which makes sure that the locations of the discontinuities remain unchanged after the enhancement has treated the coefficients and no new edges will be created.
- The function is nonlinear. Different parts of the image should be amplified with different gain. To enhance weak edges buried in the background, the enhancement function should be designed such that coefficients within the certain range are amplified with higher gain. This property of the function is shown in Figure 4.2. It clearly shows that the enhancement function elevates the values of low intensity pixels while higher values are untouched.

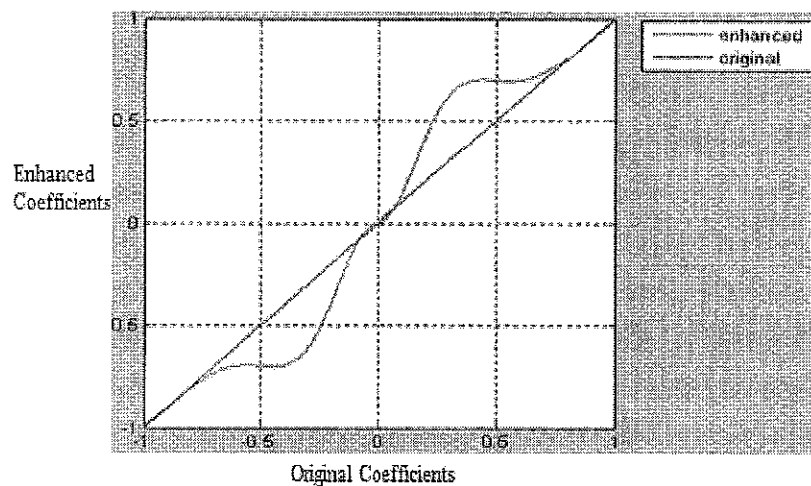


Figure 4.2 Enhanced coefficients versus original coefficients with parameters  $a=1$ ,  $b=2$ ,  $\beta=0.85$ ,  $c=4$ .

- The other most important feature of this function is its ability for denoising. This function can remove very small values in the transform domain which are always related to the image noise. The user can change the threshold of this by changing the parameter  $b$ . This is clearly shown in Figure 4.3.

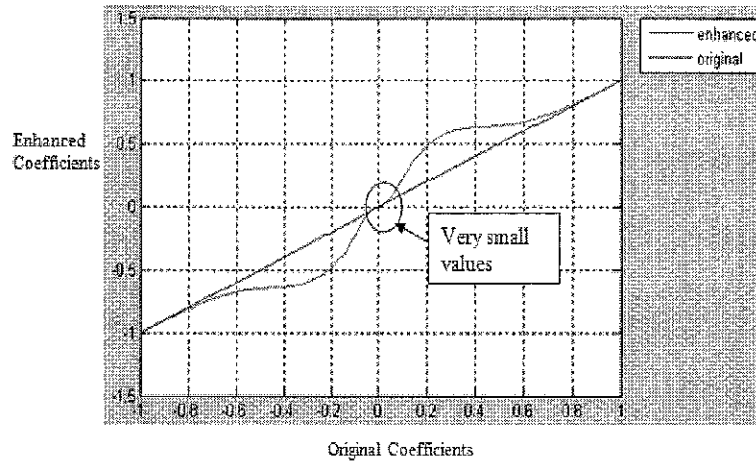


Figure 4.3 Enhanced coefficients versus original coefficients with parameters  $a=1$ ,  $b=3$ ,  $c=3$ ,  $\beta=0.85$

The over all enhancement system can be summarized by using block diagram in Figure 4.4. The normalization block helps to perform normalization by dividing each subband block with its maximum coefficient value. Thus the magnitude of every coefficient will be limited between  $[-1 \ 1]$ . The saturation image sometimes needs some modification when there exist a burst of intensity; in this case, the color of the resultant image appears to be unnatural. Thus we can linearly modify the saturation image by multiplying it by  $\alpha$ . The block is optional and after several image enhancement experiments and author's judgment the optimal range of  $\alpha$  is  $[0.6 \ 1]$ . The MR block is the multiresolution decomposition block. The block TF is for transfer or enhancement function. The hue component of the image is not touched at all because the color information should be preserved.

### 4.3 Results and Discussion

This method is tested in various types of scenes and illuminations which somehow cover all the types of images that can be found in nature but by no means exhaustive. Some of these images are taken from my personal picture database, downloaded from personal web pages, and several of them are acquired using various digital cameras mention in the

experimental setup. The graphical user interface of the developed software tool is shown in Appendix B.

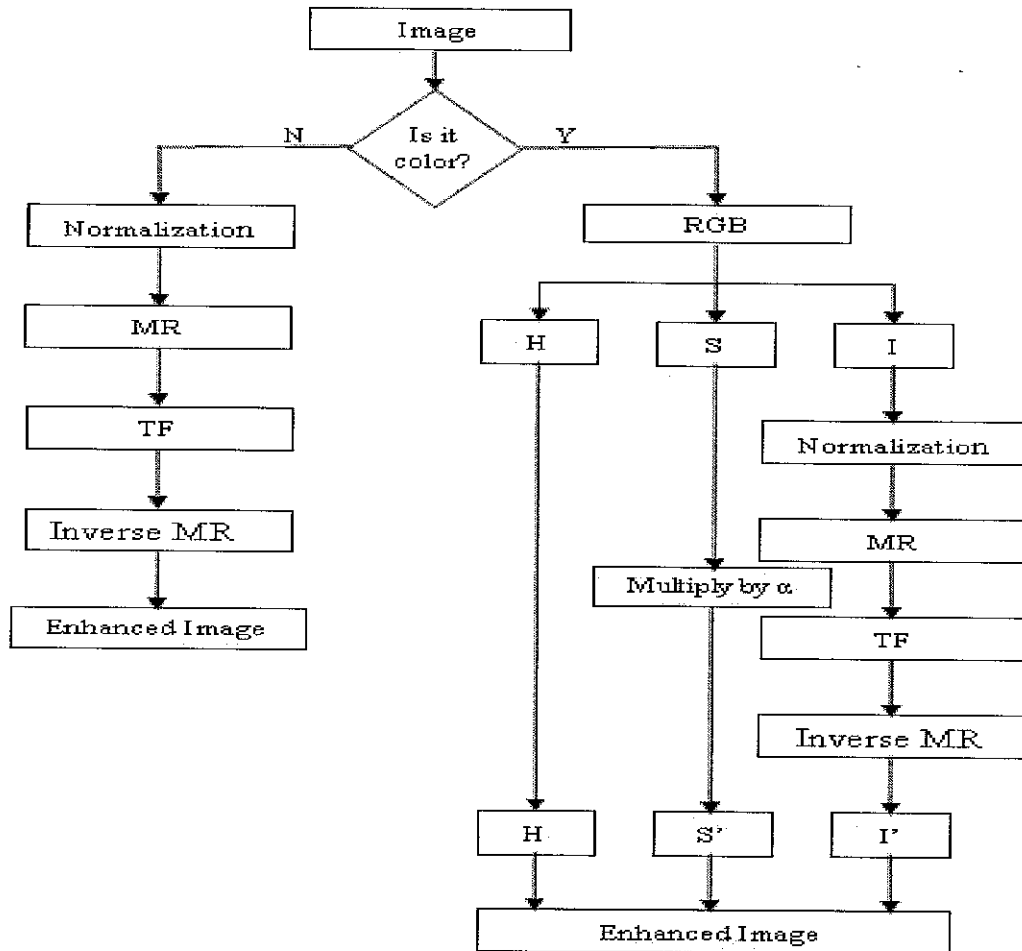


Figure 4.4. The overall general enhancement system

In this section, however we have to only show ten typical pictures as examples to demonstrate the enhanced results by our proposed method. The intensity component of the images is decomposed to 4 levels with 3, 4, 8 and 16 directions from coarse to fine decomposition levels respectively. Figure 4.5 shows the illustration of the contourlet decomposition using the peppers image. For a clear display only two decomposition levels with 4 and 8 directions is displayed.



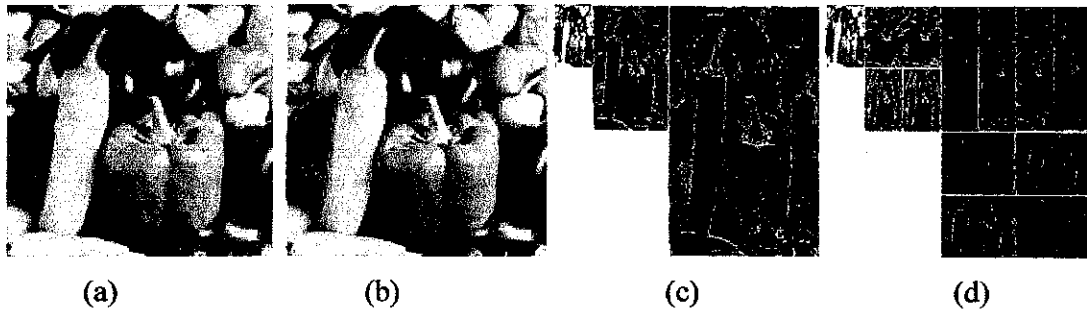


Figure 4.5 Illustration of the contourlet transform (a) true color image  
 (b) intensity image (c) pyramid transform image  
 (d) contourlet transform image

Example I: The image in Figure 4.6(a) shows a scene of a city. The image is taken in a low visible lighting condition. The image “suffers” from a low visibility. The color of the tree branches in the image appears to be dark but it is supposed to be green. Applying the enhancement system on this image gives a result which has a better clarity with no color distortion as shown in Figure 4.6 (b). The result image has a good global contrast.

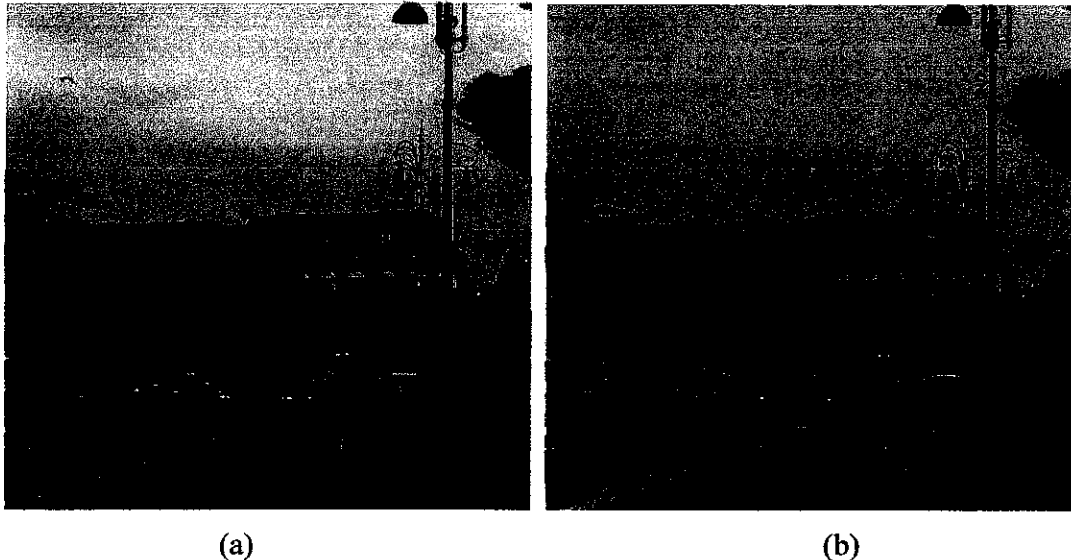


Figure 4.6 Example I with enhancement parameters  $a=1$ ,  $b=2$ ,  $\beta=0.85$ ,  $\alpha=1$ ,  $c=3$ .

Example II: Figure 4.7 (a) shows a picture of a living room. The image is taken in a non uniform lighting condition. The enhancement algorithm tries to elevate low pixel intensity values and maintains bigger values. Doing so, the dynamic range of the image is reduced which gives the enhanced image a better look without distortion. It is obvious that the optical quality of the enhanced image has improved against the degraded picture of Figure 4.7(a).

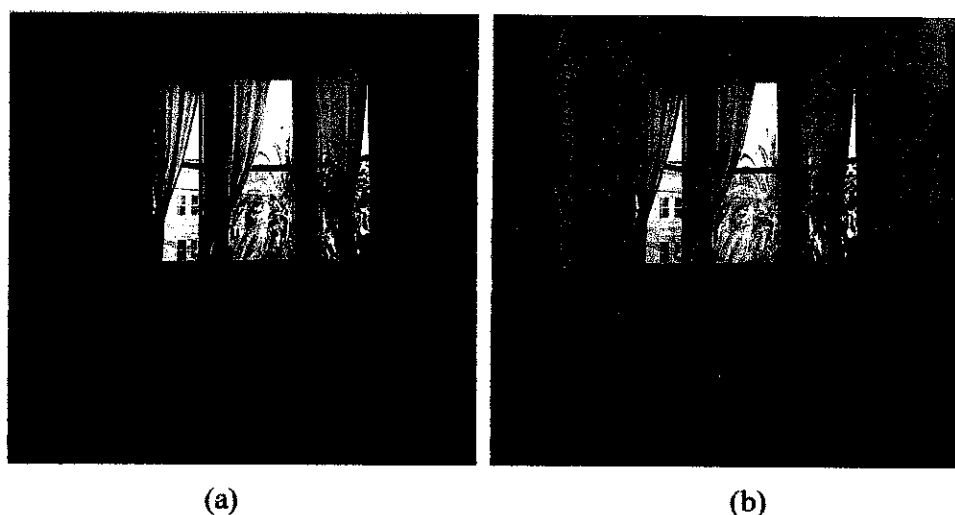


Figure 4.7 Example II, with enhancement parameters  $a=1$ ,  $b=2$ ,  $\beta = 0.85$ ,  $\alpha = 1$ ,  $c=3$ .

Example III: The image in this test, Figure 4.8(a) shows an ancient wall painting which is taken in a low but uniform lighting condition. Enhancing this image gives a more detailed as well as clear image of the original image as displayed in Figure 4.8 (b).

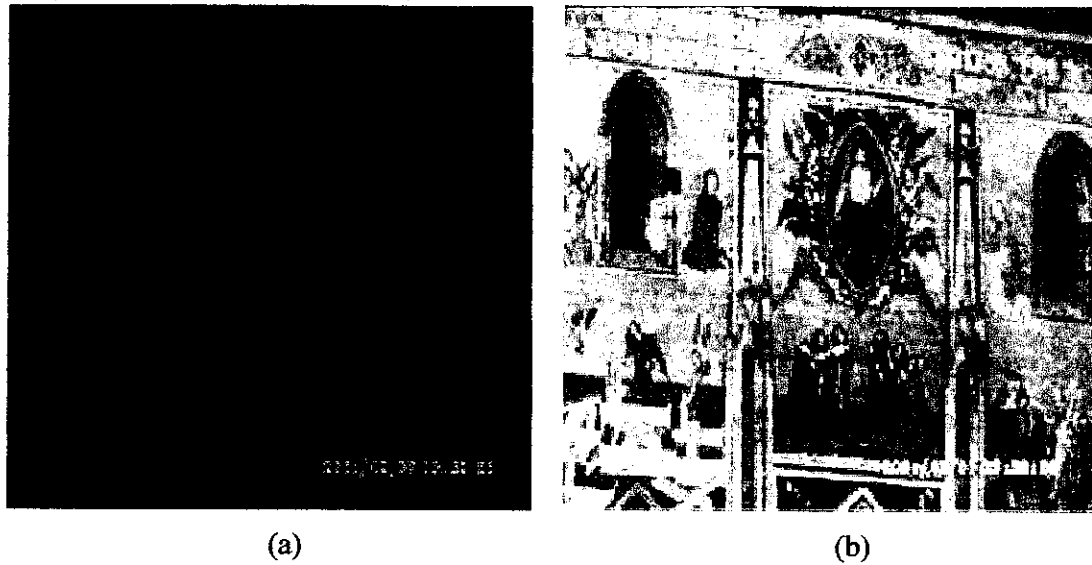


Figure 4.8 Example III with enhancement parameters  $a=1$ ,  $b=3$ ,  $\beta = 0.85$ ,  $\alpha = 0.7$ ,  $c=4$ .

Example IV: This image shows buildings and parked cars; see Figure 4.9(a). Due to the presence of the light source in the middle of the image, the cars and parts of the building are almost invisible. Applying the enhancement method on this image produces an enhanced image which displays a clear image in which both the cars and the building are visible, see Figure 4.9 (b). The enhanced image shows a slight green color distortion this is because the details of the image in the dark areas are deeply covered by dark shadows and when trying to elevate them it affects the green color spectra which is always presents a darker zone.

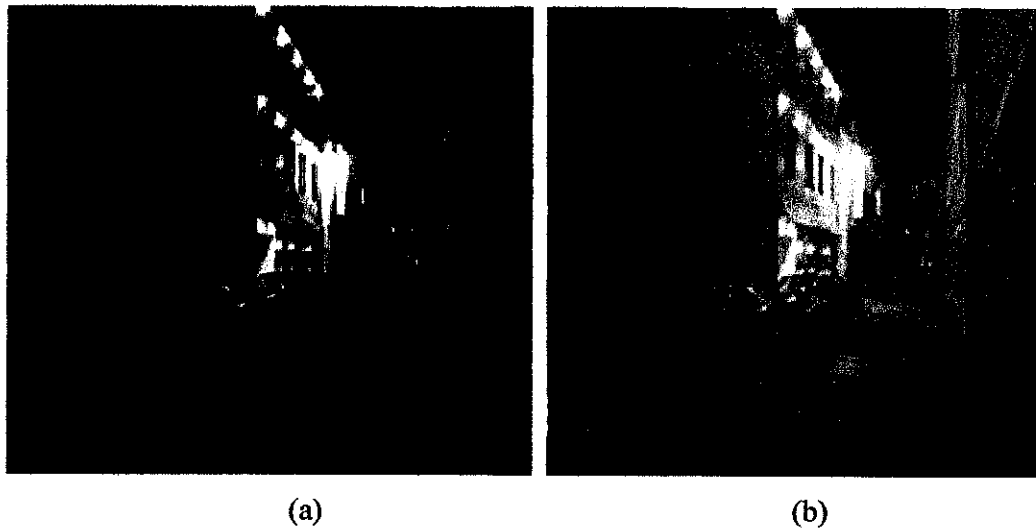


Figure 4.9 Example IV with enhancement parameters  $a=1$ ,  $b=1$ ,  $\beta = 0.85$ ,  $\alpha = 0.7$ ,  $c=4$ .

Example V: The image in Figure 4.10 (a) shows a man driving a car and two persons crossing the road. Since the image is taken at night, the persons crossing the road are almost invisible. Enhancing this image provides a more detailed scene of the road where the persons crossing the road with their dog are clearly visible. The surrounding is also clear, see Figure 4.10 (b). This shows that the enhancement method can be used as a driving assistant for night time driving or driving in low light conditions.

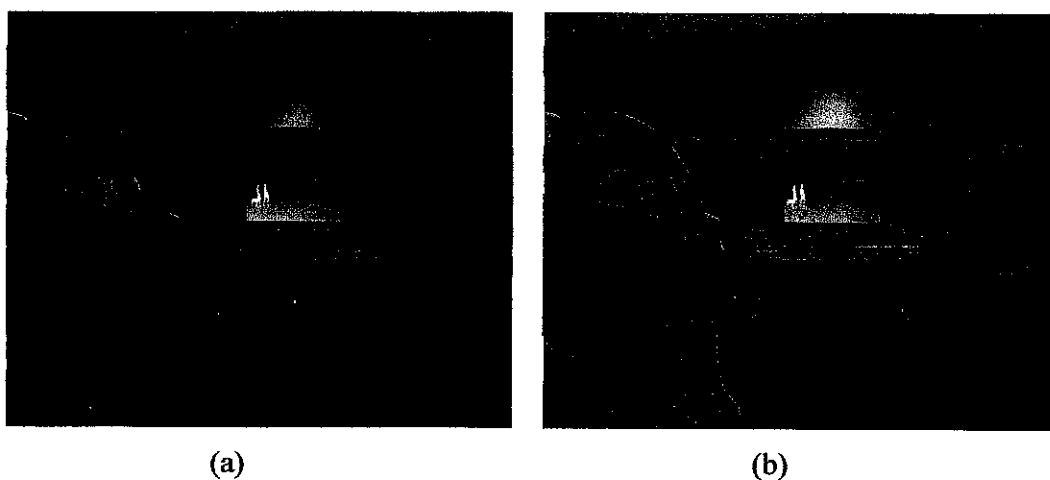


Figure 4.10 Example V, with enhancement parameters  $a=1$ ,  $b=2$ ,  $\beta = 0.85$ ,  $\alpha = 1$ ,  $c=3$ .

Example VI: The following picture shows a building, a tree, a road. The original image, Figure 4.11 (a) has a non uniform lighting condition. The presence of a light source affects the dynamic range of the image. Employing the enhancement algorithm improves the overall view of the image by reducing the dynamic range. The tree is clearly visible and the over all scene appears to be so clear and appealing. The naturalness of the image is not altered; see Figure 4.11(b).

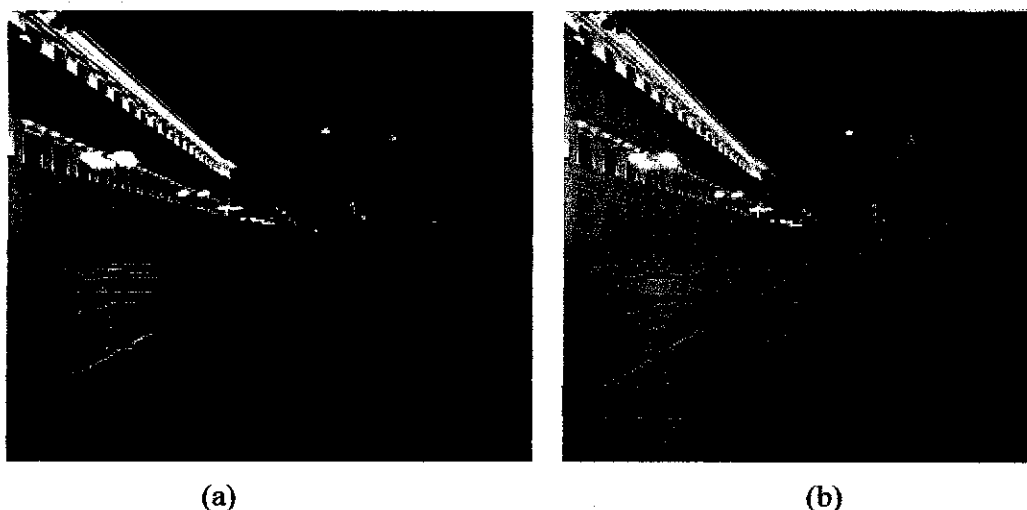


Figure 4.11 Example VI with enhancement parameters  $a=1$ ,  $b=2$ ,  $\beta = 0.85$ ,  $\alpha = 0.8$ ,  $c=3$ .

Example VII: Figure 4.12 (a) is an original nature scene with low contrast created due to a natural phenomena like fog; and the enhanced image after the enhancement algorithm is shown in Figure 4.12(b). Contrasting the original picture, we can find the green color is compensated obviously, so the whole scene is more beautiful and realistic. This image is clearly better than the original.

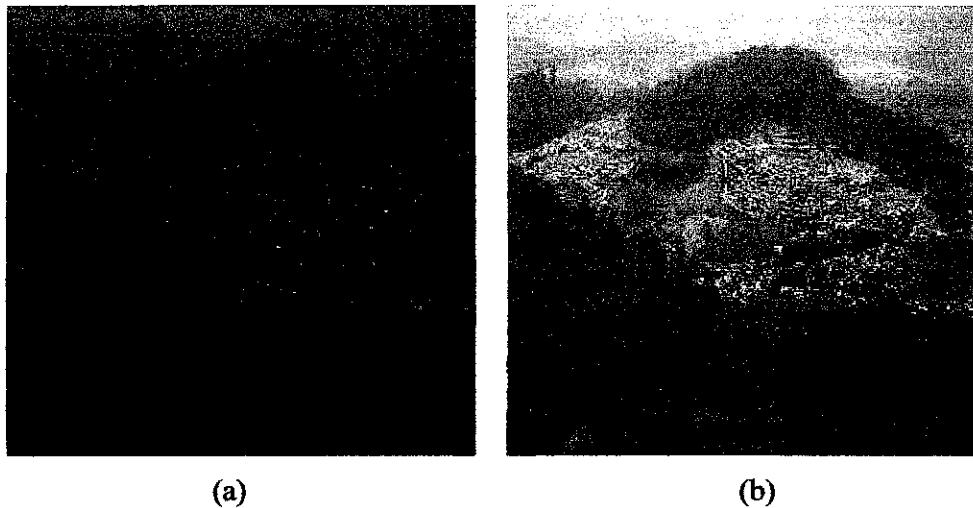


Figure 4.12 Example VII, with enhancement parameters  $a=1$ ,  $b=2$ ,  $\beta = 0.85$ ,  $\alpha = 1$ ,  $c=4$ .

Example VIII: Figure 4.13(a) illustrates the effectiveness of this image enhancement algorithm. The original image exhibits dominant dark regions created by low illumination and limited dynamic range effect. However, the enhanced image Figure 4.13(b) displays a much more visible and clear view which is similar to what human viewers perceive. The enhanced image also illustrates the effectiveness of the contrast enhancement process for producing sharp images with compressed dynamic range.

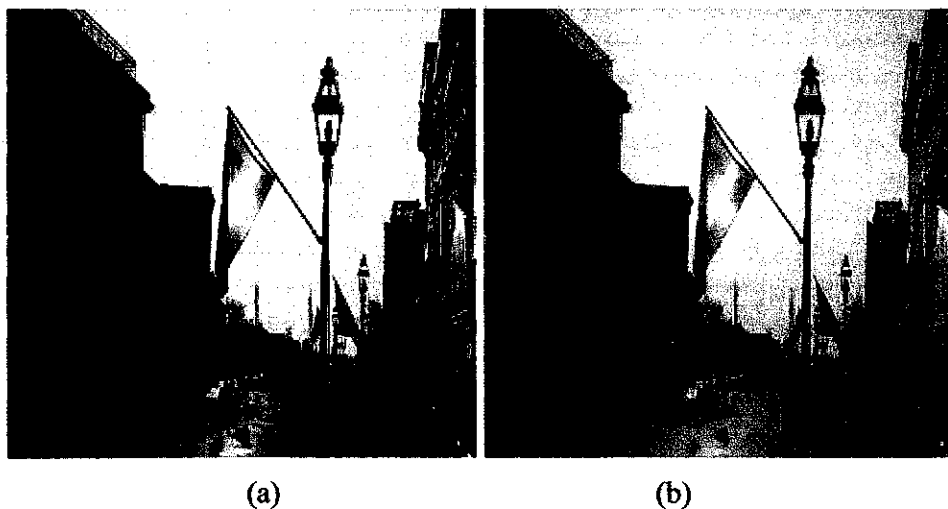


Figure 4.13 Example VIII, with enhancement parameters  $a=1$ ,  $b=2$ ,  $\beta = 0.85$ ,  $\alpha = 1$ ,  $c=3$ .

Example IX: Figure 4.14(a) shows an example of an image illuminated by the sun light at day time. The illumination is non uniform which affects the overall scene of the image. In contrast to the original picture, we find that the image enhanced by our method is clearer and brighter appears to be more natural, see Figure 4.14(b).

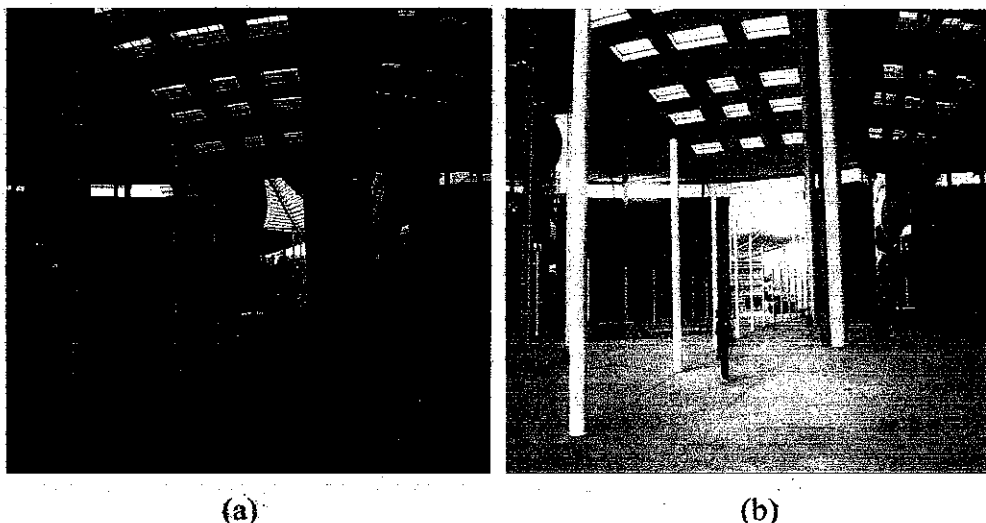


Figure 4.14 Example IX, with enhancement parameters  $a=1$ ,  $b=2$ ,  $\beta = 0.85$ ,  $\alpha = 1$ ,  $c=3$ .

Example X. The image in Figure 4.15(a) shows a man inside the vegetation. This image is taken from a typical video surveillance camera. Prior to the enhancement the man is almost invisible. Enhancing the image, see Figure 4.15(b), clearly shows that there is a man inside the bushes. This enhancement technique can be applied in security surveillance systems. However, the complete detail of the man in the enhanced image is also invisible. In this case the single image approach fails to regenerate completely hidden parts of a scene. To alleviate this problem a composite image enhancement approach is proposed. This image is used as a case study in composite image enhancement system in chapter 5.

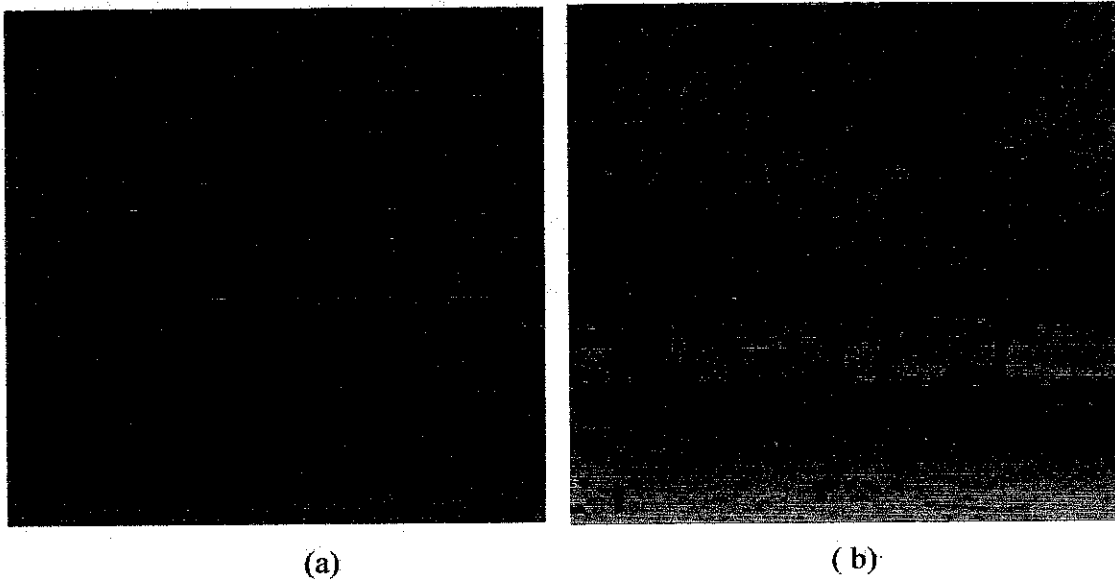


Figure 4.15 Example X, with enhancement parameters  $a=1$ ,  $b=2$ ,  $\beta = 0.85$ ,  $\alpha = 0.8$ ,  $c=3$ .

In summary, the results of these ten typical scenes images indicate that the contourlet based enhancement approach works well for images suffering from low or non uniform illumination. In the next section, we will evaluate its performance relative to other enhancement approaches, and in particular wavelet based enhancement, HE, and CLAHE.

#### 4.4 Image Quality Assessment

In image enhancement algorithms for which the resultant image is to be viewed by human beings, the only correct method of quantifying visual image quality is through subjective evaluation. In practice however, subjective evaluation is usually too inconvenient, time consuming and very expensive.

The other possible image quality metrics is a full reference objective evaluation method. This is highly dependent on the availability of an original distortion free image which the distorted image can be compared. Most existing quality assessment algorithms are grouped in this category.



Natural image signals are highly structured, their pixels exhibit strong dependencies. [Zhou Wang et al, 2004] proposed a new paradigm for quality assessment based on the hypothesis that the human visual system is highly adapted for extracting structural information. This measurement is called the structural similarity index measure (SSIM). We adopted this measure for our objective evaluation and the analysis for this method is given below.

#### 4.4.1 The SSIM Index

Consider we have two images  $x$  and  $y$ , assume one of the images to has a perfect quality, then the similarity measure can serve as a quantitative measurement of the quality of the second image. We can split this task as luminance comparison, contrast comparison and structure comparison.

For luminance comparison:

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (4.4)$$

where  $\mu$  is the estimated mean intensity  $\mu_x = \frac{1}{N} \sum_{i=1}^N x_i$  and  $\mu_y = \frac{1}{N} \sum_{i=1}^N y_i$  and  $c_1$  is defined from the dynamic range  $R$  and a constant  $k_1$ ,  $c_1 = (k_1 R)^2$ ,  $k_1$  is always defined to be  $k_1 < 1$ .

For contrast comparison:

$$c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (4.5)$$

where  $\sigma$  is the standard deviation given by  $\sigma_x = \left( \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{1/2}$  and  $\sigma_y = \left( \frac{1}{N-1} \sum_{i=1}^N (y_i - \mu_y)^2 \right)^{1/2}$  and  $c_2$  is defined from the dynamic range  $R$  and a constant  $k_2$ ,  $c_2 = (k_2 R)^2$ ,  $k_2$  is always defined to be  $k_2 \ll 1$ .

For Structural comparison:

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3} \quad (4.6)$$

where  $\sigma_{xy}$  is the correlation coefficient given by  $\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)$ , and  $c_3$  is defined from the dynamic range  $R$  and a constant  $k_3$ ,  $c_3 = (k_3 R)^2$ ,  $k_3$  is always defined to be  $k_3 \ll 1$ .

Finally combine the three comparisons,

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (4.7)$$

For simplicity let  $\alpha = \beta = \gamma = 1$  and  $c_2 = 2c_3$  then,

$$SSIM(x, y) = \frac{(2\mu_x \mu_y + c_1)(\sigma_{xy} + c_3)(2\sigma_x \sigma_y + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)(\sigma_x \sigma_y + c_3)} \quad (4.8)$$

Simplifying equation 4.8 we get:

$$SSIM(x, y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (4.9)$$

SSIM ranges from 0 to 1. Zero is for completely different images and 1 for perfect match. Results nearer to one are acceptable. For the experiment  $k_1=0.01$  and  $k_2=0.03$  are used.

Image signals are non-stationary thus it is a good practice to measure over local regions rather than measuring through the entire image. In this case we propose a sliding window ( $w$ ) approach. The window starts from the top-left corner of the images and ends at the bottom-right. Then the total index measure will be found by averaging the local measures. The window can have a size of  $4 \times 4$  or  $8 \times 8$ . The windowed structural similarity index measure is then given by:

$$WSSIM(x, y) = \frac{1}{|W|} \sum_{w \in W} SSIM(x, y|w) \quad (4.10)$$

where  $W$  is the family of all windows and  $|W|$  is the cardinality of  $W$ .

#### 4.4.2 Evaluation Setup

The performance evaluation image set contains 31 natural images having different contrast levels and varying illumination. This set contains set of images that are artificially fabricated from perfect image by reducing the illumination, blurring and adding Gaussian noise. Hence we can use the windowed structural similarity index measure (WSSIM) and the percentage signal to noise ratio (PSNR) between the original image and the enhanced images. To assess the speed of the algorithm the computational time is also measured. This is the total amount of CPU time which the processor takes to process the whole enhancement procedure.

Figure 4.16(a) shows an example image for the objective evaluation. The image contains a chair and four binders on it taken using a digital webcamera in a room where the windows are opened and the light is turned on. This image serves as a reference image to measure the WSSIM and PSNR. Figure 4.16(b) is the same picture taken when the

windows are closed and the light is turned off, still there is a little illumination but not enough to take a clear image. This is the image to be enhanced using different methods. Comparing results obtained when the test image is processed by HE (Figure 4.16(c)), CLAHE (Figure 4.16(d)), Wavelet (Figure 4.16(e)) and Contourlet methods (Figure 4.16(f)); the contourlet method gives us a greater WSSIM index and PSNR as shown in Table 4.1. But it takes the maximum time to process. This is because the contourlet transform is defined with double filter banks which actually take more time to generate the coefficients.

This is obvious that the contourlet image has more brightness which is considerably different to that of the dark test image. The smooth shapes of the binders and the chair are maintained. HE and Wavelet contain artifacts that are not originally in the image. Even though the HE method appears to be brighter; the color of the original image has considerably changed which is not wanted. In the CLAHE method the enhanced image appears to be dark and its performance is inferior as compared with the contourlet method.

Table 4.1. Objective evaluation result for the example image from Figure 4.16

Method	WSSIM	PSNR	Computational Time (in seconds)
HE	0.9861	19.7333	0.447898
CLAHE	0.9989	22.8724	0.541503
Wavelet	0.9991	23.4461	0.819482
Contourlet	0.9992	25.8193	0.931454
Test Image	0.3644	12.0913	---

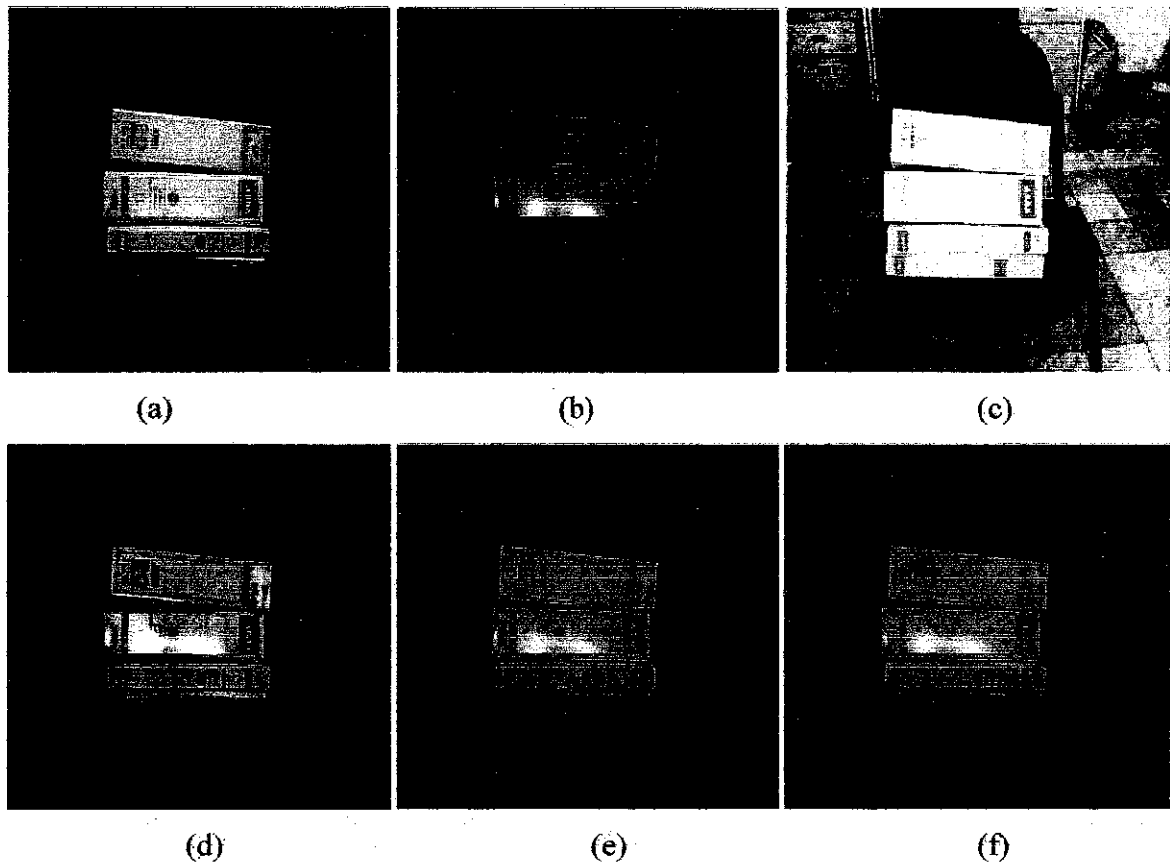


Figure 4.16 Objective evaluation test image (a) Perfect reference image (b) test image (c) HE Method (d) CLAHE (e) wavelet enhanced image (f) contourlet enhanced image

To visually compare the color intensity profile, we can display each spectral band profile across the image. The graphs in Figure 4.17 clearly show that the contourlet method has the nearest profile to the perfect image. This tells us that the contourlet method does not introduce any artificial distortions. On the other hand the profiles of HE and CLAHE reveal there is a distortion.

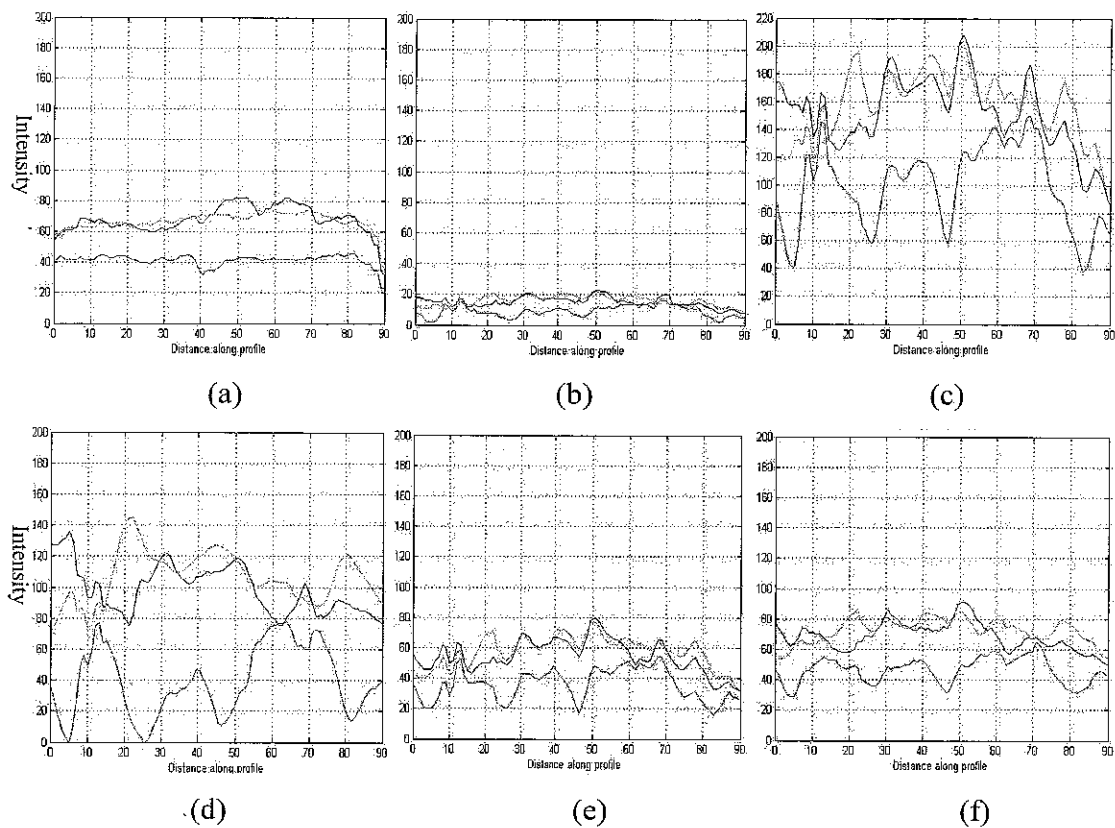


Figure 4.17 RGB profile of the result from the Figure 4.16 (a) Profile of the perfect image (b) Profile of the test image (c) profile of the HE Method (d) profile of the CLAHE Method (e) Profile of the Wavelet Method (f) Profile of the Contourlet Method

Table 4.2 shows the average windowed structural similarity index (AWSSIM), average peak signal to noise ratio (APSNR) and average computation time it takes to enhance of all the 31 images used in the objective evaluation of our algorithm. The result in this table also shows that the contourlet method is superior both in AWSSIM and APSNR and inferior in terms of computational complexity.

Table 4.2: Objective Evaluation for 31 images

Method	AWSSIM	APSNR	Average Computational Time ( In seconds)
HE	0.8661	20.4409	1.112065
CLAHE	0.8911	23.8722	1.327968
Wavelet	0.9701	27.5491	1.781384
Contourlet	0.9749	27.6623	2.172756
Test Image	0.3831	11.9843	--

Each value of the AWSSIM of the 31 test images is tabulated and drawn using a line chart in Figure 4.18. From this figure we can see that the contourlet method is most of the time superior and the HE is mostly inferior.

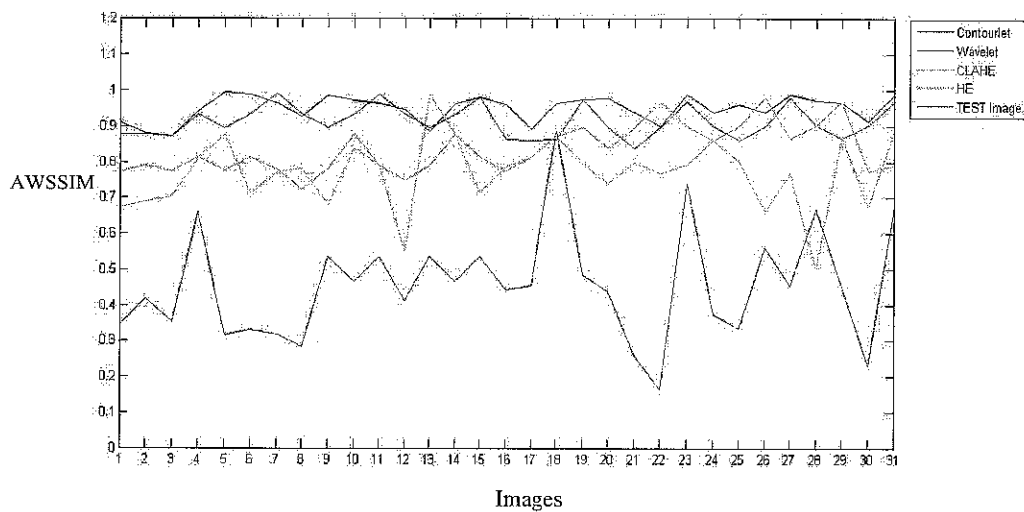


Figure 4.18 Comparison of all the four methods for all 31 images.

#### 4.4.3 Comparison with the Existing Image Enhancement Technique based on Contourlet Transform

Our approach is compared with another existing contrast enhancement method proposed by P. Feng [P. Feng et al, 2007]. In this work they proposed a retinal image enhancement technique based on contourlet transform. They used the Stark's enhancement function [Jean-Luc Starck et al, 2003]. The main reason for the choice of contourlet transform is for its better performance of representing edges and textures of natural images, which means better representation of lesions and blood vessels of a retinal image.

For a fair comparison we used the same image taken from a standard retinal image source—the Utrecht DRIVE database. According to their paper [P. Feng et al, 2007] the green channel of colored retinal images gives highest contrast between vessels and background, this is channel is chosen to be used in the enhancement procedure. The enhancement results are of the P. Feng method and our method is displayed in Figure 4.19. When we look at the enhanced images with the two methods, there is no considerable difference which can be detected visually. The P. Feng et al method uses the enhancement function shown in equation 2.9, with parameters  $m = 30$ ,  $c = 3$ ,  $s = 0.6$ , and  $p = 0.5$ . For our method the parameters used are  $a=1$ ,  $b=2$ ,  $\beta = 0.85$ ,  $\alpha = 1$ ,  $c=3$ . It is not possible to perform the objective performance analysis using the WSSIM as there is no ground truth. But observing on the result we can roughly say that the performance of both methods is good as the blood vessels are clearly visible in both the enhanced images.



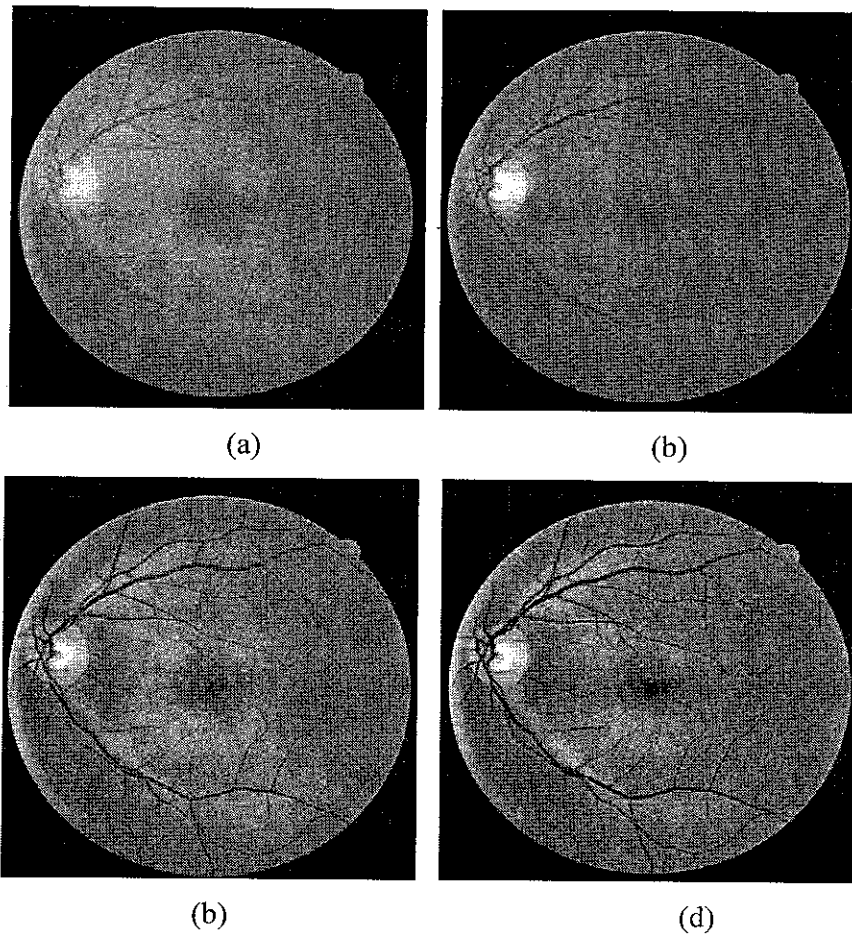


Figure 4.19 Comparison of our method with the P. Feng contrast enhancement method  
(a) True color retinal image (b) green band (c) our method (d) P. Feng method

Table 4.3 shows the computational time taken for the enhancement algorithms. Since the P. Feng method uses a lot of comparison statements in the enhancement function it takes more computational time, on the other hand our method uses only one nonlinear function thus it has less time. The test was performed on two images with dimension 512x512 and 256 x256.

Table 4.3 Computational time for retinal images

Test Images	Computational time ( in seconds)	
	P. Feng et al Method	Our Method
Image 1	3.421353	3.218208
Image 2	0.705747	0.600453

#### 4.5 Color Space Selection

A number of color image enhancement algorithms have been developed in the past. Some of these algorithms have shown that they are very efficient and powerful. They employ one of the several kinds of color space representation techniques existing in the literature. Most of the algorithms are set up for use with any of the three component color spaces. However, there is little substantive information about which color space should we use for an optimal enhancement result.

The objective of this section is to determine whether there is a difference in performance when we employ each color space for a given enhancement algorithm and to adapt that color space for our enhancement technique. Ten most common color spaces are examined to assess this result. Both numerical and psychophysical techniques are used for evaluation purposes.

##### 4.5.1 Comparison based on Enhancement Performance

The windowed structural similarity index measurement technique is used to assess the enhancement results for each color space. Thirty one images were used in this experiment. Each image is artificially distorted from the original perfect image by using Gaussian blurring, illumination reduction, and adding noise. The original image is considered as a perfect image and it serves as a comparison for the performance of each enhancement. A sample result is shown in Figure 4.20.

WSSIM is measured for each of the thirty one images used with three different enhancement algorithms and the average value is displayed in Table 4.4. According to this experiment the HSI color space gives a maximum index measure in all the three enhancement algorithms, the contourlet method being the best among them. This implies that for image enhancement algorithms it is a better practice if we use the HSI color space rather than other spaces.

Table 4.4 WASSIM for the test images using in three different enhancement algorithms

Color Space	CLAHE	WAVELET	CONTOURLET
	WASSIM	WASSIM	WASSIM
HSV	0.8905	0.9715	0.9706
HSI	0.8916	0.9740	0.9743
LAB	0.8890	0.9618	0.9637
LUV	0.8790	0.9638	0.9630
LCH	0.8810	0.9678	0.9537
YCbCr	0.8913	0.9551	0.9641
YDbDr	0.8881	0.9641	0.9541
YUV	0.8613	0.9681	0.9686
RGB	0.8467	0.9218	0.9425
XYZ	0.8648	0.9398	0.9451
Test Image	0.3831	0.3831	0.3831

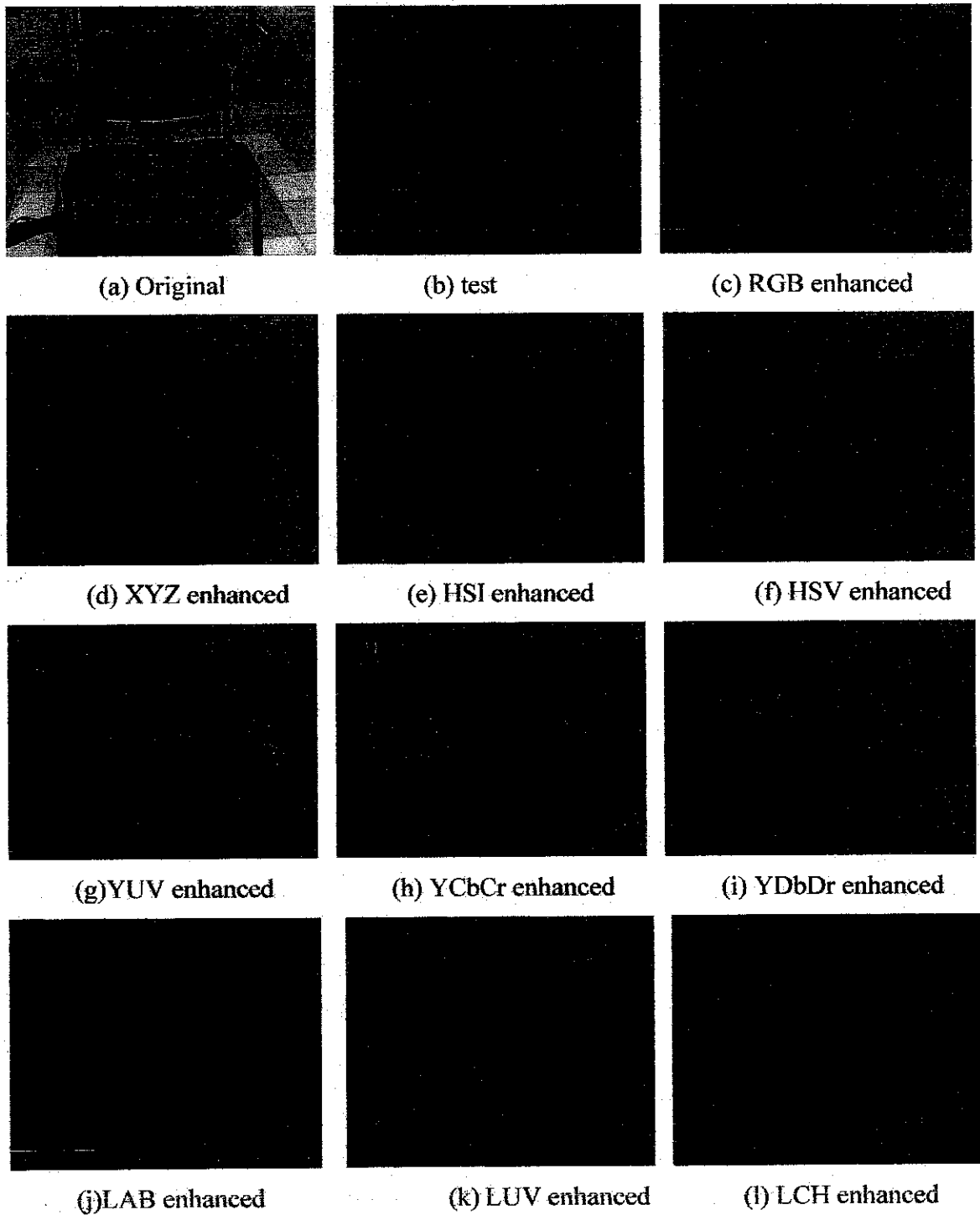


Figure 4.20 Enhancement of the Chair image in different color spaces based on the contourlet transform method

The image shown in Figure 4.20 (a) is considered as a perfect image. The image in Figure 4.19(b) is the same image artificially modified by severely reducing the ambient illumination. This image is one of the thirty one test images used to assess the color spaces. Enhancement using the contourlet transform method is performed in each of the ten color spaces. Figure 4.20 (c), (d), (e), (f), (g), (h), (i), (j), (k) and (l) shows all the enhancement results for color spaces RGB, XYZ, HSI, HSV, YUV, YCbCr, YDbDr, LAB, LUV and LCH respectively. The final result indicates that RGB is the worst color space. The HSI is the best. HSV is the second best. The other color spaces are just intermediate. Thus in conclusion, the HSI color space is a good choice for color image enhancement application. In the next chapter (Chapter five) the HSI color space is used as the color representation for fusing composite images.

## CHAPTER FIVE: COMPOSITE IMAGE ENHANCEMENT

### 5.1 Introduction

Composite image enhancement can be broadly defined as the process of combining multiple input images into a smaller collection of images, usually a single one, which contains the salient information from the inputs, in order to enable a good understanding of the scene, in terms of semantic interpretation. It is always referred to as an image fusion process.

Fusion process can take place at different levels of information representation. A common categorization is to distinguish between pixel and region level. Image fusion at pixel level amounts to integration of low-level information, in most cases physical measurements such as pixel intensity. It generates a composite image in which each pixel is determined from a set of corresponding pixels in the various sources.

Fusion at region level requires first the segmentation procedures to extract features contained in the various input sources. Those features can be identified by characteristics such as size, shape, contrast and texture. The fusion is thus based on those extracted features and enables the detection of useful features with higher confidence. Fusion at region level allows the combination of information at the highest level of abstraction. The input images are usually processed individually for information extraction and classification.

Fusion can also be classified as a multi-focal and multi-modal depending on the number and types of sensors involved. The later uses multiple cameras while the former uses a single camera with different focal points. In this thesis a generic pixel based image fusion system which works well for both multi-focal and multi-modal systems is proposed.

The aim of image fusion is to integrate complementary and redundant information from multiple images to create a composite image that contains a better description of the scene than any of the individual source images. Therefore the composite image should offer the following basic advantages:

- By integrating information, image fusion can reduce dimensionality. This results in a more efficient storage and faster interpretation of the output.
- By using redundant information, image fusion may improve accuracy as well as reliability, and
- By using complementary information, image fusion may improve interpretation capabilities with respect to subsequent tasks. This leads to more accurate data, increased utility and robust performance.

Considering the objectives of image fusion and its potential advantages, some generic requirements can be imposed on the algorithm:

- it should not discard any salient information contained in any of the input images;
- it should not introduce any artifacts or inconsistencies which can distract or mislead a human observer or any subsequent image processing steps;
- it must be reliable, robust and, as much as possible, tolerant to imperfections such as noise or misregistration.

Image fusion is an intrinsic behavior of all animals. Human beings use two eyes (multi-focal system) to clearly perceive an object and precisely determine the distance, speed and location of the object in the scene. One of the most famous examples of multimodal image fusion in living organisms is the visual system of rattlesnakes. These vipers possess organs which are sensitive to thermal radiation. The infrared signals provided by these organs are combined by bimodal neurons with the visual information obtained from

the eyes. Inspired by this real-world example, several researchers have tried to imitate it using multi-sensor image fusion [Gema Piella, 2003, H. Wang, 2004].

Another biologically inspired fusion method is the approach based on multiresolution (MR) decomposition. It is motivated by the fact that the human visual system is primarily sensitive to local contrast changes, i.e. edges, and MR decompositions provide a convenient spatial-scale localization of such local changes. The basic strategy of a generic MR fusion scheme is to use specific fusion rules to construct a composite MR representation from the MR representations of the different input sources. The composite image is then obtained by performing the inverse decomposition process.

Henceforth, we confine our discussion to MR image fusion approaches. In particular, we focus on MR fusion schemes (working at pixel level) where the output is a single composite image which is constructed primarily for display on a computer monitor.

## 5.2 Pixel Based Image Fusion

Today, most image fusion applications employ pixel fusion methods [Gema Piella, 2003, Yin Chen, 2007]. The advantage of pixel fusion is that the images used contain the original information. Furthermore, the algorithms are rather easy to implement and time efficient. As we observed before, an important preprocessing step in pixel based fusion methods is image registration, which ensures that the data at each source is referring to the same physical structures. As a reminder, it will be assumed that all source images have been registered or we will use existing primitive alignment techniques.

Figure 5.1 shows a detailed block diagram of the fusion scheme. The framework contains eight most important modules. A description of each to the building blocks is given next with example.



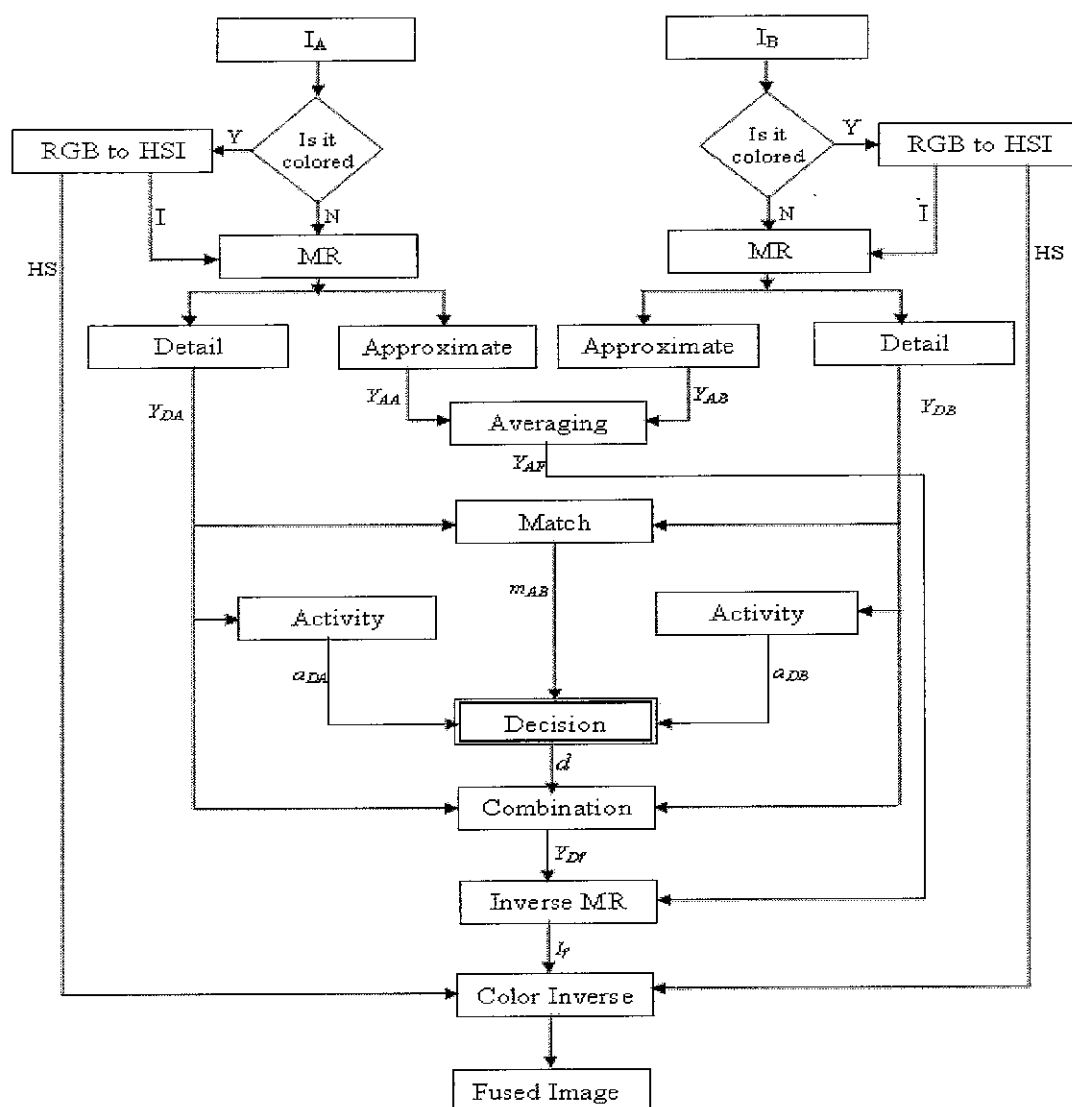


Figure 5.1 Generalized pixel based MR image fusion scheme with two input images and a composite image output

### 5.2.1 MR Analysis

This block computes an MR decomposition of the input sources. Given two input images as shown in Figure 5.2, the MR decomposition is first done on the intensity component of each of the images using the contourlet transform. Generally we have used a 4 level

decomposition and detail coefficients at each level are decomposed to 3, 4, 8 and 16 decomposition directions starting from coarse to fine approximation levels. This example image is taken in a typical area surveillance scenario for monitoring an outdoor scene in a bad visibility condition. A man is walking through the scene behind the bushes. He is hardly visible in the visual image while the night vision camera captures him clearly but not the details of the vegetation and color.

Thus appropriately combining these two images will generate a complete composite image.

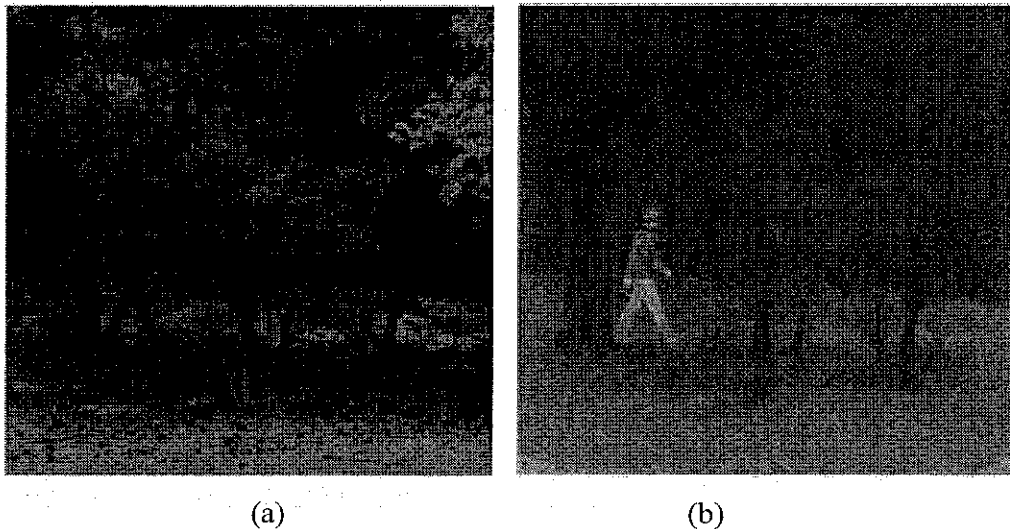


Figure 5.2 Sample Images a) visual image b) night vision image

Because of their different physical meanings, the approximate and detail coefficients are treated by the combination algorithm in a different fashion. Figure 5.3 shows the contourlet transform coefficients of the intensity components of the images from Figure 5.2 with 3 levels (for display purposes only) and 3, 4, 8 directions from coarse to fine approximations respectively.

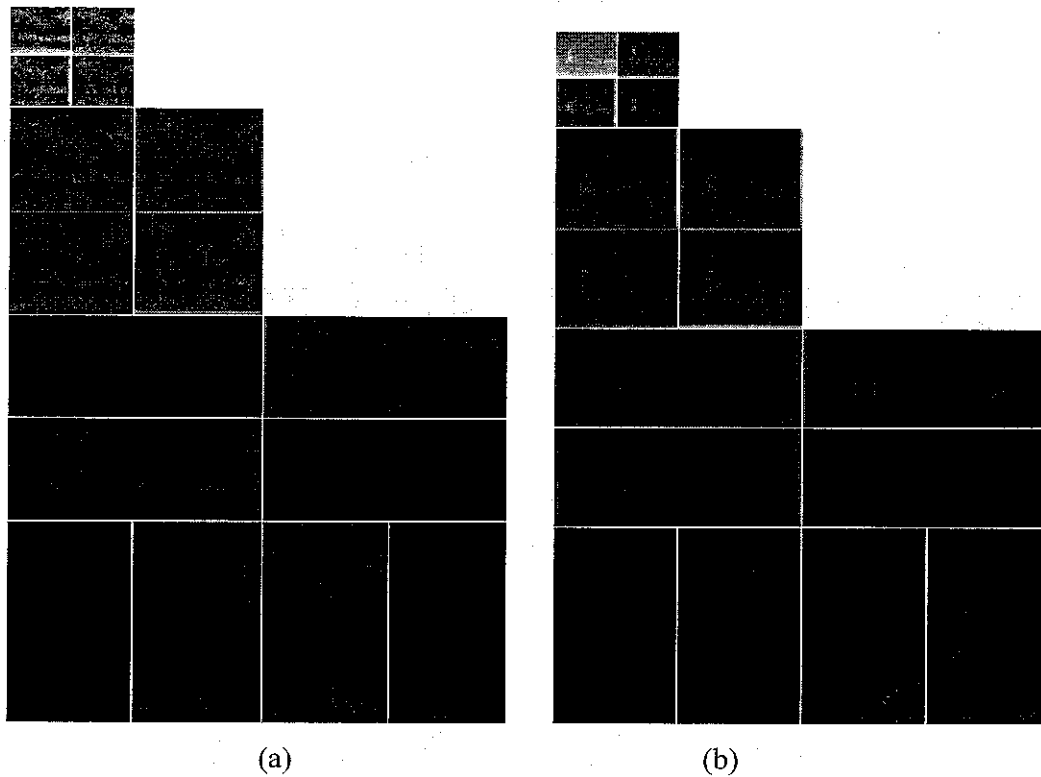


Figure 5.3 Contourlet transform coefficients of the source images, a) visual image b) night vision image

### 5.2.2 Color Transformation

Usually the visual images carry color information which is very important in semantic interpretation and scene understanding. Color images are represented by three spectral bands, RGB. Performing the analysis on each spectral band is a huge amount of task but instead the color image is first converted to an appropriate color space, HSI (Hue, Saturation, and Intensity), which effectively separates the color information and the intensity. We perform our analysis on the intensity image only. A detail explanation about color spaces and why we chose the HSI color space is given in chapter 5. Figure 5.4 shows the H, S, and I components of the visual image of Figure 5.2(a).

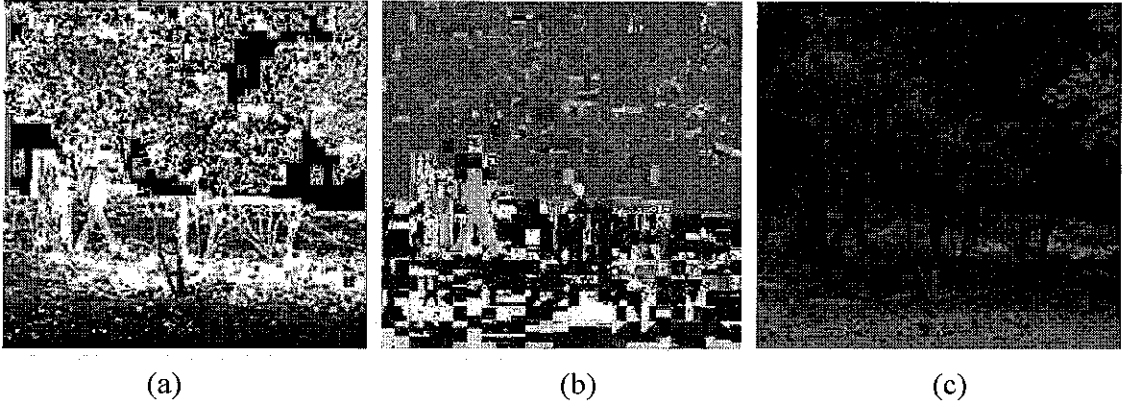


Figure 5.4 HSI decomposition of the visual image (a) hue (b) saturation (c) intensity

### 5.2.3 Approximate Coefficients

The approximate images designated by  $Y_{AA}$  and  $Y_{AB}$  represent a coarse representation of the original image  $I_A$  and  $I_B$  respectively. They inherit most of its properties such as the mean intensity and some coarse texture information. Thus, coefficients  $Y_{AA}$  and  $Y_{AB}$  with high magnitudes do not necessarily correspond with salient features. Therefore, the composite image  $Y_{AF}$  is a simple averaging of the input images as given in equation 5.1.

$$Y_{AF} = \frac{Y_{AA} + Y_{AB}}{2} \quad (5.1)$$

However, there is a technical reason behind this simple averaging, which is based on the assumption that the source images contain additive Gaussian noise and the decomposition level is high enough, a minimum of four levels in our case, thus important image features have already been captured by detail coefficients. Thus, the approximate images contain mainly noise and averaging them reduces the variance of the noise while ensuring appropriate mean intensity.

### 5.2.4 Detail Coefficients

For detail images the relevant perceptual information relates to the edge information that is present in each of the detail coefficients  $Y_{DA}$  and  $Y_{DB}$ . Detail coefficients having large absolute values correspond to sharp intensity changes and hence to salient features in the image such as edges, lines and region boundaries. Thus the fusion is done as follows based on the works of Burt and Kolczynski. Match and activity measures will be performed on the coefficients. From these two measures we will construct a decision matrix which is a key for the combination of the images. The final composite image is found by taking the inverse contourlet transform of the combination output.

### 5.2.5 Activity Measure

Activity measure is the measure of the degree of saliency of each coefficient in the detail coefficients  $Y_{DA}$  and  $Y_{DB}$ . (i.e. its importance to image fusion at hand). The saliency should increase when features are in focus (for multi focal systems) or it should give emphasis on the contrast difference (for multi modal systems). In MR decomposition systems the contrast information can be attributed by the magnitude of the detail frequency components. Higher amplitude implies a good contrast. The fact that human visual system is highly sensitive to local contrast changes (edges) lets us to compute the activity as a local energy measure [P. J. Burt P. J, and Kolczynski, 1993] as given in equation 5.2.

$$a_{Di}^{(l,k)} = \sum_{\Delta n \in W^{(l,k)}} \left| Y_{Di}^{(l,k)}(x + \Delta n, y + \Delta n) \right|^2 \quad (5.2)$$

where  $l$  is the decomposition level,  $k$  is subband direction in level  $l$ ,  $i$  is input image in this case images A and B, and  $W$  is a finite local window of size  $3 \times 3$ .

The activity measures of the source images (Figure 5.2) can be visualized as a gray scale image in Figure 5.5. Higher gray values pixel indicates a higher saliency of that pixel. The man in Figure 5.5 (b) shows a higher gray level thus its saliency is very high.

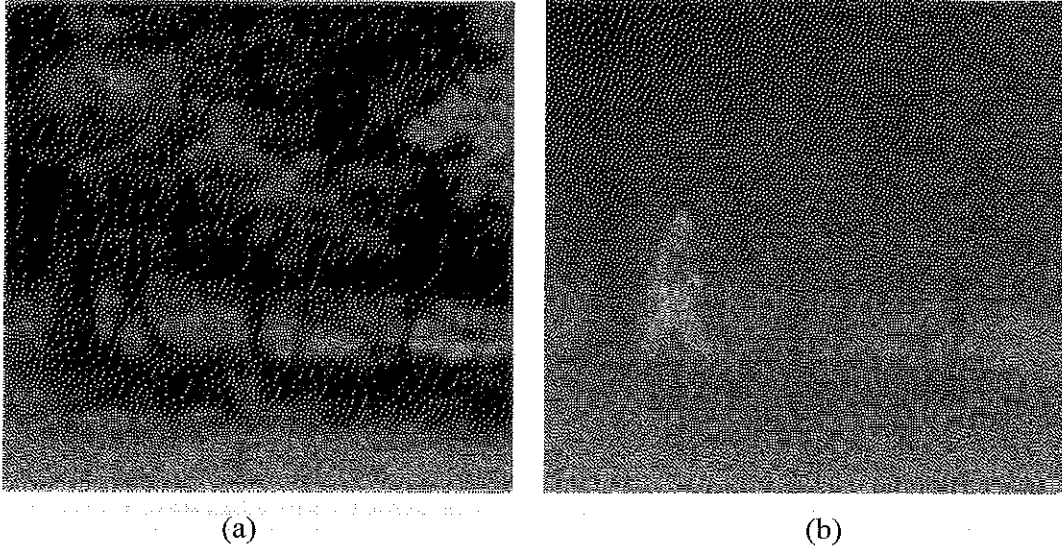


Figure 5.5 Visualization of the Activity measures (a) Visual Image (b) night vision image

### 5.2.6 Match Measure

Match measure is used to quantify the degree of similarity between the source images. Precisely  $m_{AB}^{(l,k)}(x,y)$  reflects the resemblance between the detail coefficients  $Y_{DA}$  and  $Y_{DB}$  for decomposition level  $l$  and direction  $k$ . The match measure tells where the sources are different and to what extent they differ. We can use this information to appropriately combine them. In order to measure it correctly we should also consider the neighborhood coefficient values. Thus it is defined as a normalized correlation average over neighborhood of the samples [P. J. Burt P. J, and Kolczynski, 1993] as shown in equation 5.3 where  $W^{(l,k)}$  being window of size  $3 \times 3$ .

$$m_{AB}^{(l,k)}(x,y) = \frac{2 \sum_{\Delta n \in W^{(l,k)}} Y_{DA}^{(l,k)}(x+\Delta n, y+\Delta n) Y_{DB}^{(l,k)}(x+\Delta n, y+\Delta n)}{a_{DA}^{(l,k)}(x,y) + a_{DB}^{(l,k)}(x,y)} \quad (5.3)$$

The match measure of the sample test images (Figure 5.2) can be visualized using the Figure 5.6 as a gray scale image. The values of match measure have been linearly scaled to the range [0 255] for display. The gray scale value at each pixel location represents the similarity of the neighborhood pixels. Higher pixel values indicate higher degree of resemblance.

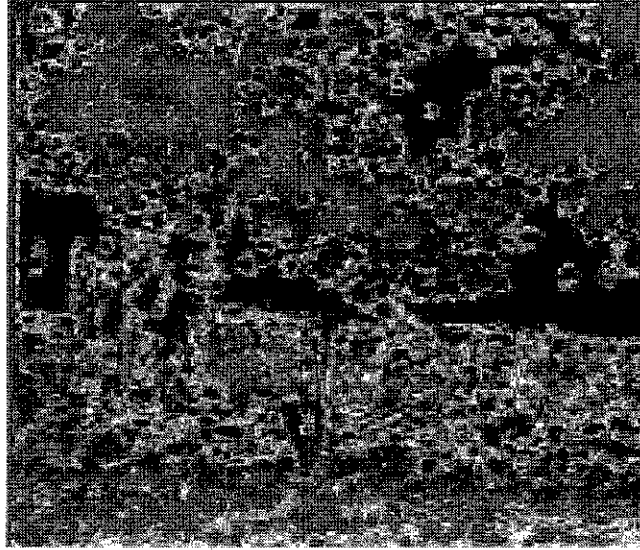


Figure 5.6 Visualization of the match measure

### 5.2.7 Decision Map

Decision map is the core of the combination algorithm. Its output governs the actual combination of the coefficients of the various sources. It controls the weights to be assigned to each source coefficients. The conventional approach is to assign a weight which is directly proportional to the activity measure. But this has a contrast reduction effect in images which have opposite contrast; which actually assumes that at a point only one of the source images is a valid choice [Gema Piella, 2003, Yin Chen ,2007]. The

ultimate solution is to apply a selection rule which considers the match measure and the activity measure together. At points where the source images are different the combination selects the most salient component. At location where they are similar, average the source components. The averaging provides stability and reduces noise. In this case it is computed as shown in equation 5.4 where  $T$  is decision threshold,  $T$  is defined by the user and it is always chosen to be a value nearer to one,  $T=0.85$  for this experiment.

$$d^{(l,k)}(x,y) = \begin{cases} T & \text{if } m_{AB}^{(l,k)}(x,y) \leq T \text{ and } a_{DA}^{(l,k)}(x,y) > a_{DB}^{(l,k)}(x,y) \\ 1-T & \text{if } m_{AB}^{(l,k)}(x,y) \leq T \text{ and } a_{DA}^{(l,k)}(x,y) \leq a_{DB}^{(l,k)}(x,y) \\ \frac{1}{2} + \frac{1}{2} \left( \frac{1 - m_{AB}^{(l,k)}(x,y)}{1-T} \right) & \text{if } m_{AB}^{(l,k)}(x,y) > T \text{ and } a_{DA}^{(l,k)}(x,y) > a_{DB}^{(l,k)}(x,y) \\ \frac{1}{2} - \frac{1}{2} \left( \frac{1 - m_{AB}^{(l,k)}(x,y)}{1-T} \right) & \text{if } m_{AB}^{(l,k)}(x,y) > T \text{ and } a_{DA}^{(l,k)}(x,y) \leq a_{DB}^{(l,k)}(x,y) \end{cases} \quad (5.4)$$

where  $d^{(l,k)}(x,y)$  is the weight of the first image (visual image), and  $(1 - d^{(l,k)}(x,y))$  is the weight of the second candidate image (infrared image).

Visualization of the decision map of the test images (Figure 5.2) is shown in Figure 5.7. The decision map of the infrared image (Figure 5.7(b)) contains more highlighted pixels than those in the visual image decision map (Figure 5.7(a)). This tells us that more information is merged in the fused image from the infrared image than that of the visual image. This is because the visual image contains more dark areas.



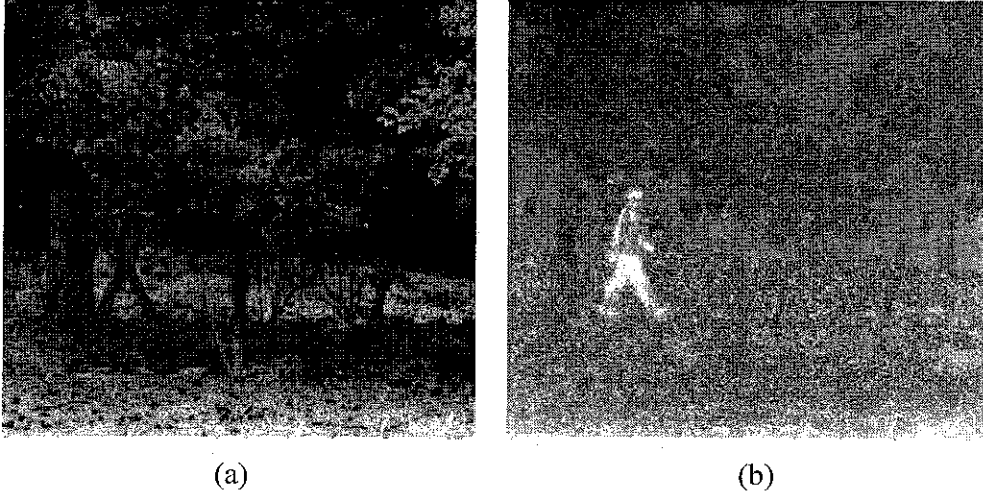


Figure 5.7 Decision Map (a) Visual Image (b) night vision image

### 5.2.8 Combination Map

Combination map describes the actual combination of the transformed detail coefficients of the source images. In this thesis we use the weighted average where the weights are determined from the decision map (equation 5.4) using equation 5.5.

$$Y_{Df}^{(l,k)}(x, y) = d^{(l,k)}(x, y)Y_{DA}^{(l,k)}(x, y) + (1 - d^{(l,k)}(x, y))Y_{DB}^{(l,k)}(x, y) \quad (5.5)$$

Now we merge fused approximate coefficients  $Y_{Af}^{(l,k)}$  and fused detail coefficients  $Y_{Df}^{(l,k)}$  to produce the complete composite coefficients  $Y_f$  which will be feed to the MR synthesis block.

### 5.2.9 MR Synthesis

Finally, the composite image is obtained by applying the inverse contourlet transform on the composite MR decomposition  $Y_f$ ,

$$I_f = \Phi^{-1}(Y_f) \quad (5.6)$$

Using  $I_f$  as the new intensity image and the hue (H) and saturation (S) components from the original visual image, a color transformation is done which converts the HSI image to RGB image for display purposes. The resultant image is shown Figure 5.8. Directional information introduced by the contourlet transform yields the best description of the background vegetation and the man inside it. The composite image is more complete than any of the source images and looks natural and the noise level is minimal.

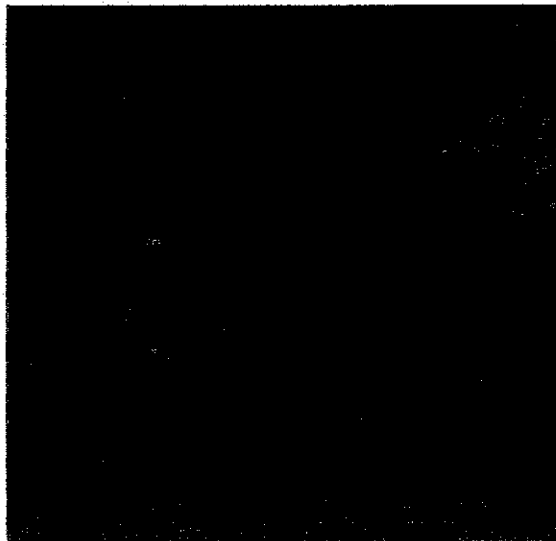


Figure 5.8 Composite Image

### 5.3 Results and Discussion

The system is tested for several kinds of images for performance evaluation. Some of typical results are displayed in the next section. This system can be applied in context enhancement, multiple focus images for faster interpretation, medical imaging, military applications, fire fighting and rescue operations, night time driving and security surveillance. A discussion of some typical image pairs is also given. The graphical user interface of the developed software tool is shown in Appendix B.

### 5.3.1 Driving at Night

The man by the side of the car is almost invisible in the visual image of Figure 5.9 (a) while we can clearly see him in the night vision image 5.9(b) in spite of the fact that there is no color information. The composite image shown in Figure 5.9(c) is precise and contains all relevant information for clear visual perception.

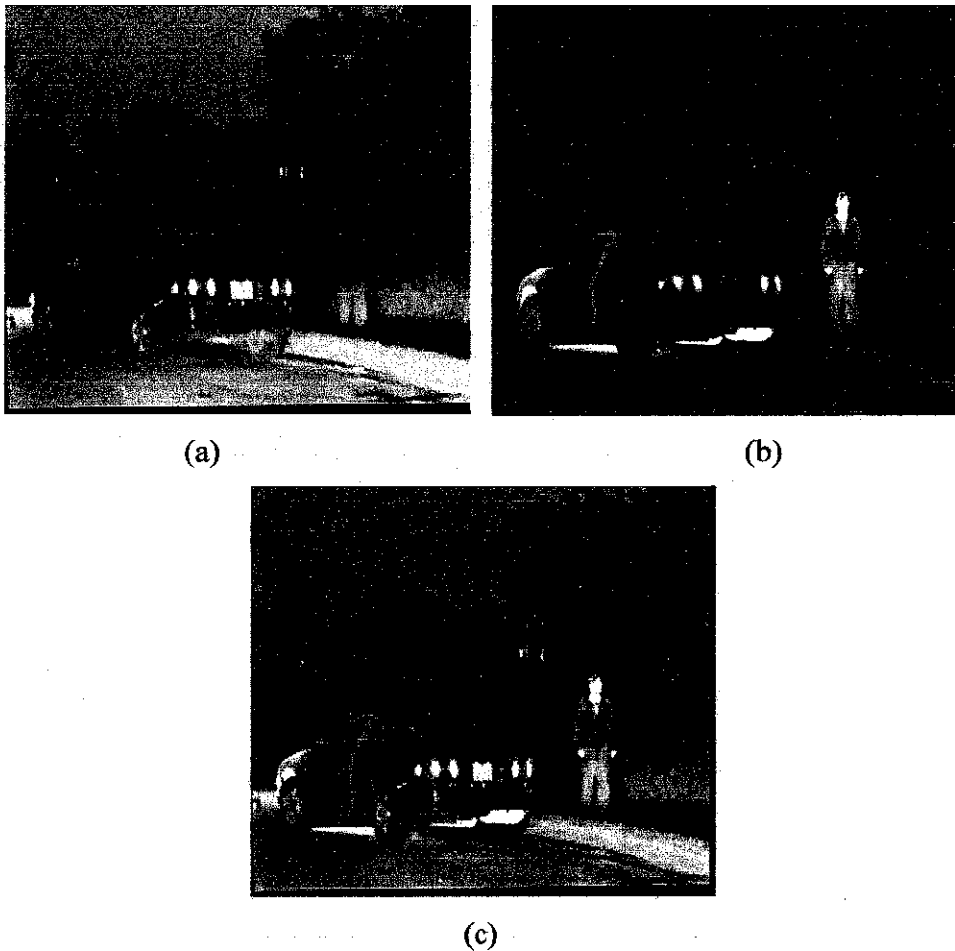


Figure 5.9 Night time driving (a) Visual camera (b) night vision camera  
(c) Composite image

### 5.3.2 Fire Fighting and Rescue Operations

Image fusion can be applied in fire fighting and rescue operation. The image in Figure 5.10 (a) is taken by a visual camera. It shows all the color information of the houses and the smoke. The second image in Figure 5.10(b) is taken by infrared camera [source: OCTEC Limited]. It reveals the details of the houses and three men. Using either of the images alone is misleading for a fire fighters group to successfully perform the rescue operation so we fuse the two images and the composite image in Figure 5.10(c) shows all the salient information.

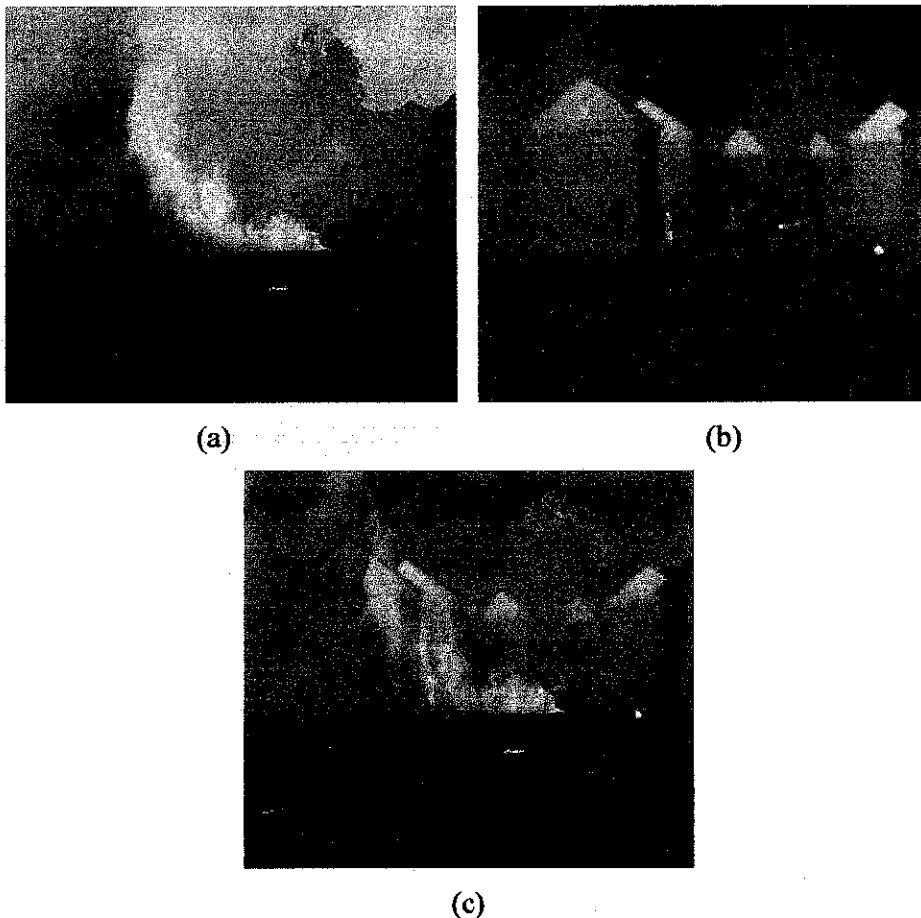


Figure 5.10 Fire fighting and rescue operation (a) visual camera (b) infrared camera  
(c) composite image

### 5.3.3 Multi-Focal Gray Image

The image in Figure 5.11 shows the application of this method in multifocal images. This image is kindly provided by [Oliver Rockinger, 2008]. Figure 5.11(a) is a near focused image where the smaller watch is clearly visible while the larger is not. Figure 5.11(b) is a far focused image. Combining these two images will produce an image where both the watches are clearly displayed. All the details are displayed very well.

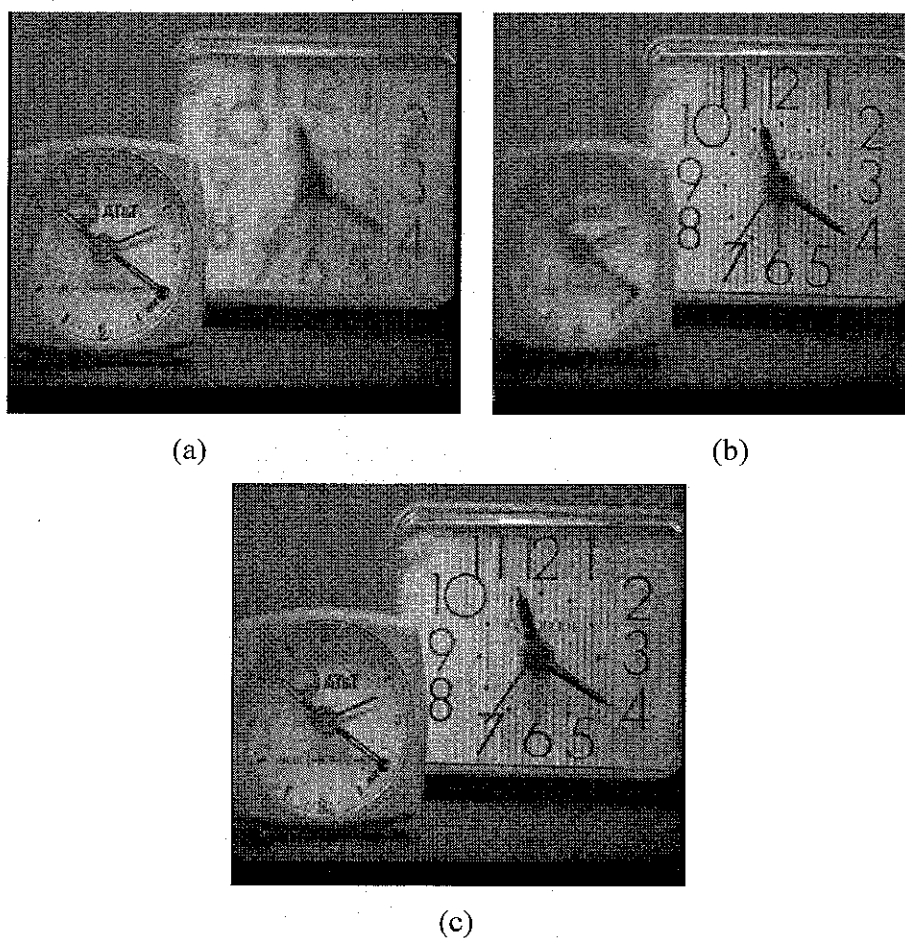


Figure 5.11 Multi-focal gray scale images (a) Near focused (b) far focused  
(c) composite image

### 5.3.4 Medical Imaging

Fusion of multimodal images can be very useful for clinical application such as diagnosis, modeling of human body or treatment planning. This image is kindly provided by [Oliver Rockinger, 2008]. The example in Figure 5.12 shows the usage of fusion in radiotherapy and skull surgery. Normal and pathological soft tissues are better visualized by magnetic resonance imaging (MRI) Figure 5.12(a); while the structure of tissue bone is better visualized by computed tomography (CT) in Figure 5.12(b). The composite image depicted in Figure 5.12(c) not only provides the salient information from both images simultaneously; but also reveals the relative position of soft tissue with respect to the bone structure.

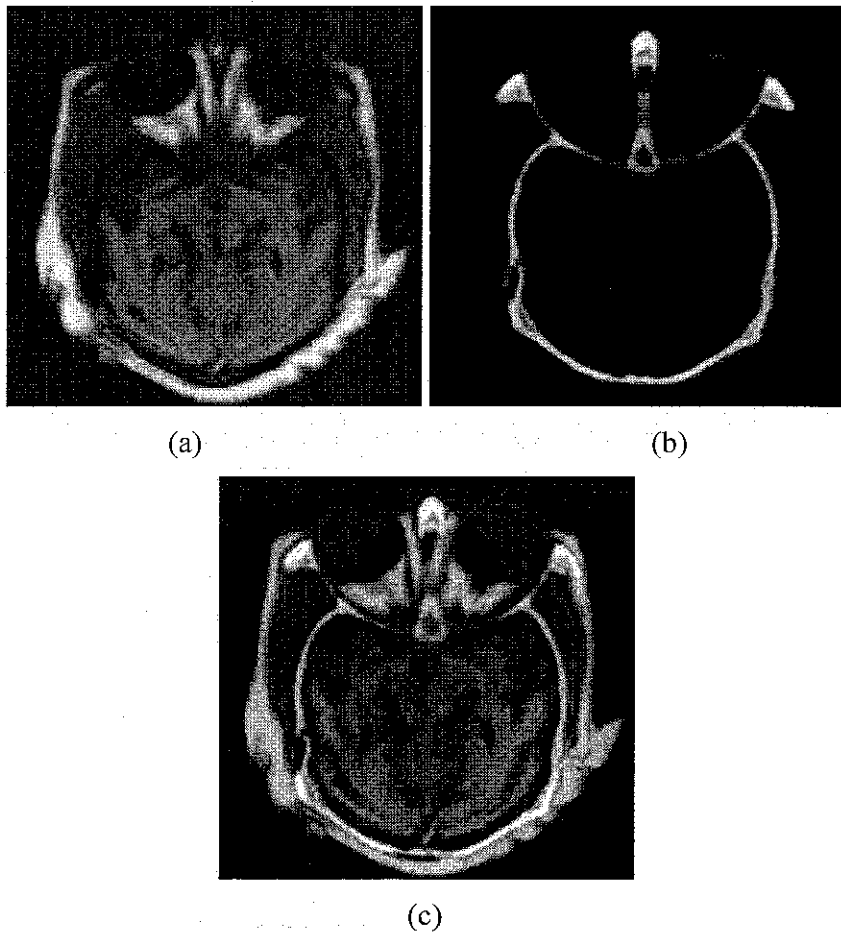


Figure 5.12 Medical imaging (a) MRI image (b) CT image (c) Composite image

### 5.3.5 Military Application

Consider the source images in Figure 5.13 (a) and (b) below depicting the same scene. In the visual image, Figure 5.13(b) it is hard to distinguish the person in the camouflage from the background. This person is clearly observable in the infrared image of Figure 5.13(a). In contrast, the easily discernible background in the visual image, such as the fence, is nearly imperceptible in the IR image. The fused image represents the overall scene better than any of the two individual images.

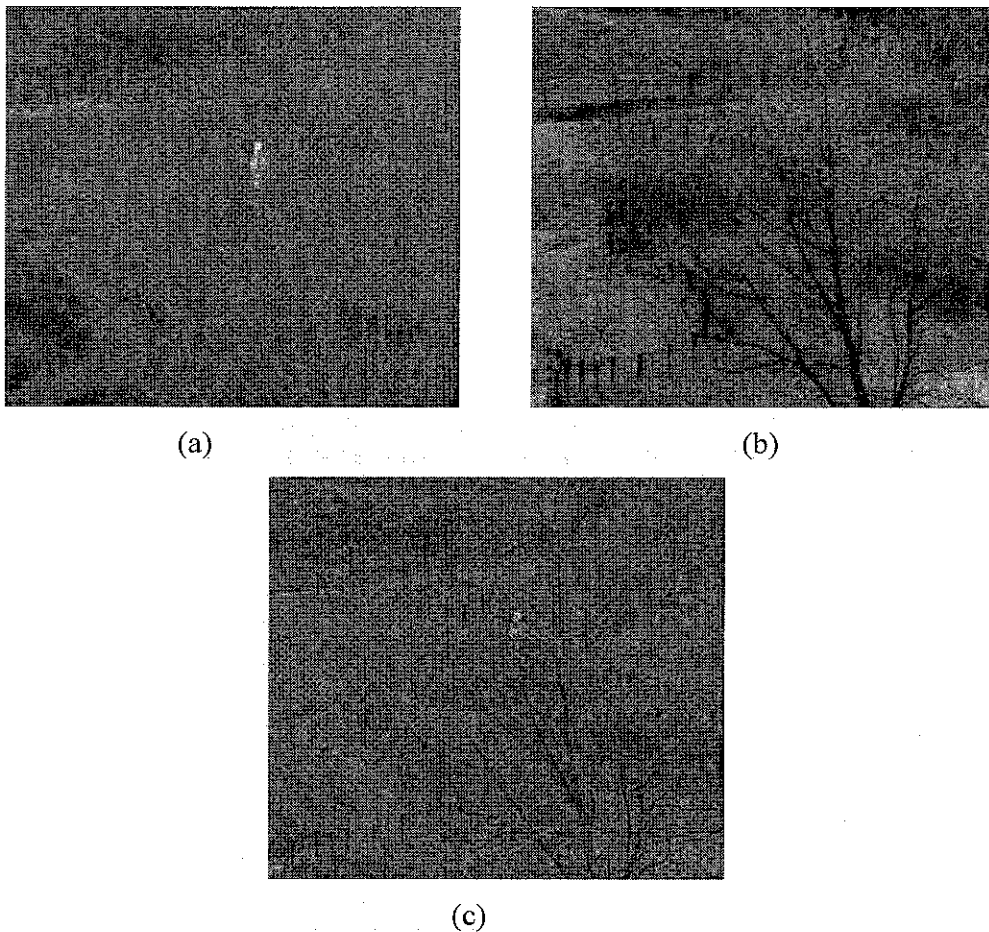


Figure 5.13 Military application (a) Infrared image (b) visual image (c) composite image

### 5.3.6 Red Eye Removal

When capturing photographs in low lighting environments it is a common practice to use the flash mode of the cameras. Doing so will give us a brighter image which has a greater signal-to-noise ratio that reveals the details of the scene that would always be hidden in a photograph acquired under ambient illumination without the flash. However using the flash may introduce artificial artifacts such as a red eye. The optimal solution for this problem is to take a pair of images one with flash and the other with no flash. These two images can be combined using this algorithm and the result is encouraging. We slightly modify the algorithm as follows. In the color transformation phase the intensity components will be combined in the same manner but we should always take the saturation (S) component from the picture taken with flash and the hue (H) component from the image taken with no flash. Doing so will retain all the brightness of the image while removing the red eye.

The sample result in Figure 5.14 shows the application of the fusion method in red eye removal of a girl. Figure 5.14(a) shows a picture taken with the flash is on and has red eye. The picture in Figure 5.14(b) is the same image taken using no flash mode (this image is kindly provided by G. Petschnigg, M. Agrawala). This image suffers from low illumination but there is no red eye. Combining these two images creates a more complete image without the red eye as shown in Figure 5.14(c). The advancement of digital photography has made it possible to quickly and automatically capture pair of image like this thus fusion will be an optimal solution.



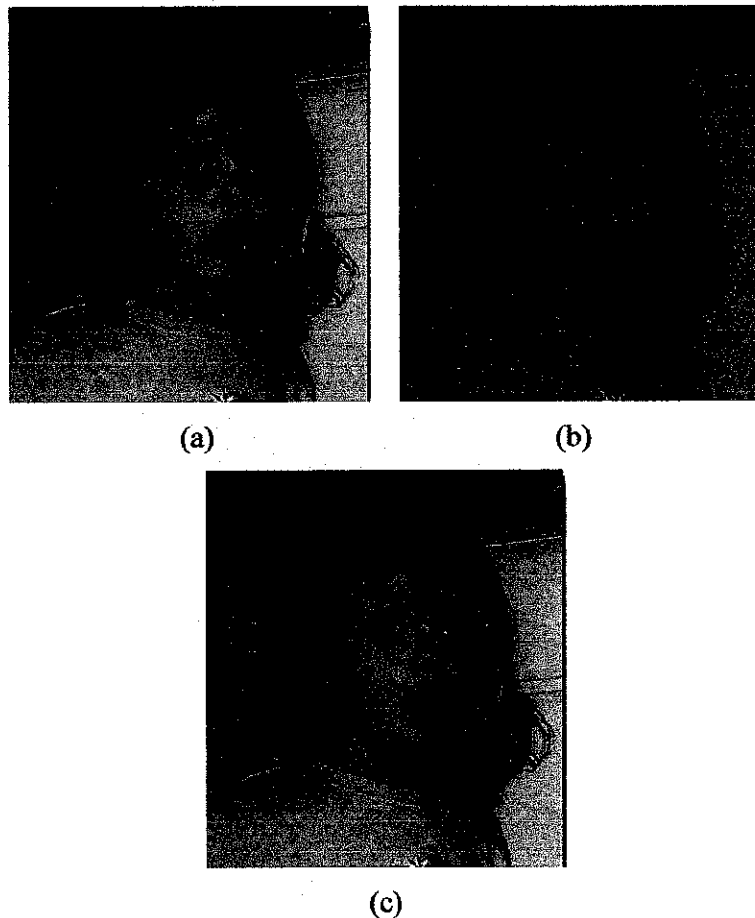


Figure 5.14 Red eye removal (a) Image taken when the flash is on (b) image taken when the flash is off (c) fused image

#### 5.4 Performance Evaluation

Here we propose a full reference objective performance assessment technique. Due to the lack of a clearly defined ground-truth the test images here have been created artificially by blurring the original images by using Gaussian blurring technique. Hence here we can use the windowed structural similarity index measure (WSSIM) which is discussed in chapter 4 section 4.4.1 and the peak signal to noise ratio (PSNR) between the original image and the generated fused images. The computational time is also computed for each test image.

The first pair of test images used are complementary pairs of a man as shown in Figure 5.15(b) and (c.) They are created by using Gaussian blurring of radius 10 from the original image shown in Figure 5.15(a). The fused images by contourlet transform wavelet transform and simple averaging are shown in Figure 5.15(d), (e) and (f) respectively.

Table 5.1 compares the quality of these composite images using the windowed structural similarity index (WSSIM) and PSNR between the original and the generated fused images. The numerical results in the Table 5.1 show that the contourlet result is way superior to the wavelet method. All the methods gave good indexes as the test images provided are easy to combine. However, the contourlet transform method gives the maximum computation time. This is mainly because the construction of the contourlet transform consists of double filter banks which take higher time to generate the coefficients.

Table 5.1 Performance evaluation result of the test image on Figure 5.15

<b>Fusion Method</b>	<b>WSSIM</b>	<b>PSNR</b>	<b>Computational Time (in seconds)</b>
<b>Contourlet Method</b>	<b>0.99649</b>	<b>34.7126</b>	<b>4.171588</b>
<b>Wavelet Method</b>	<b>0.99612</b>	<b>31.9957</b>	<b>4.051072</b>
<b>Simple Averaging</b>	<b>0.93573</b>	<b>26.5733</b>	<b>1.288362</b>

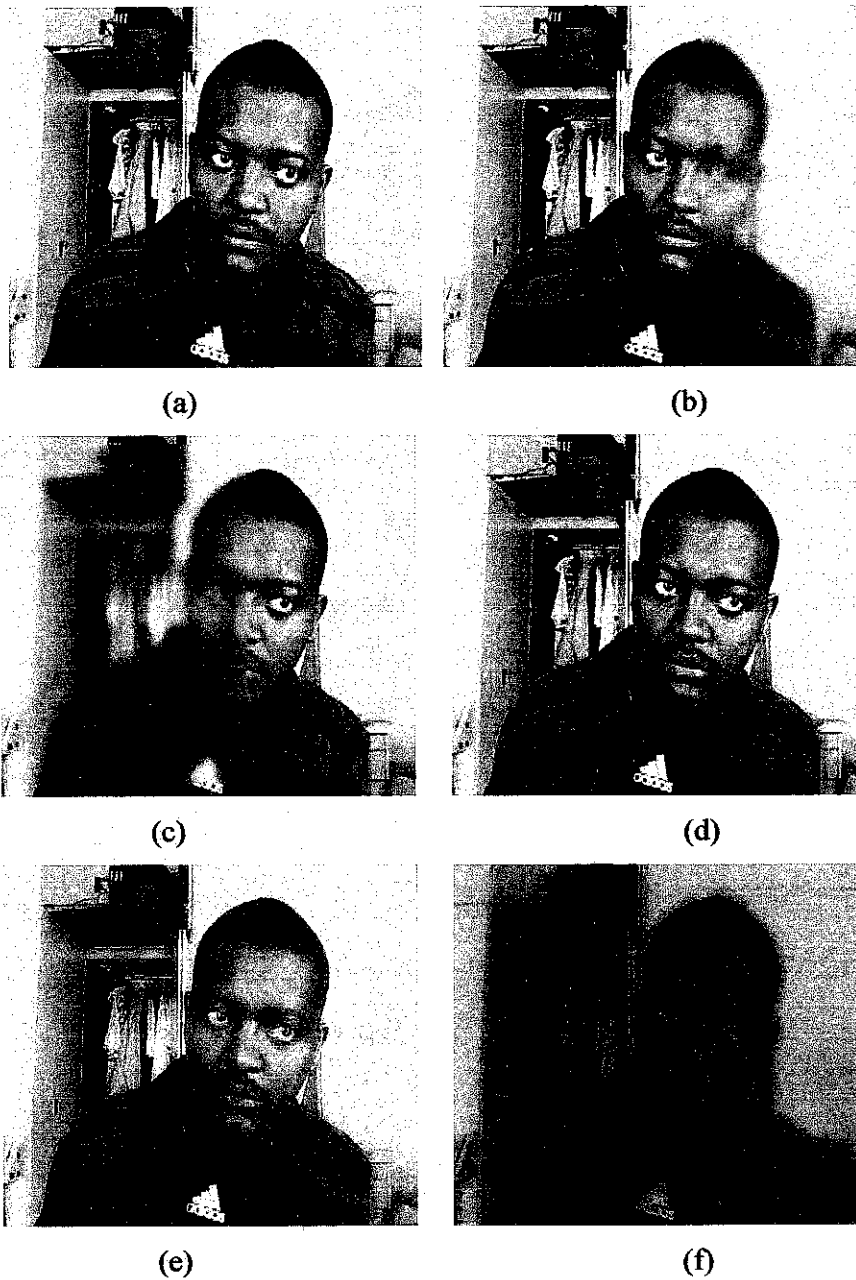


Figure 5.15 Objective evaluation test image (a) original reference image (b) right side blurred image (c) left side blurred image (d) contourlet method (e) wavelet method (f) simple averaging

The second objective test image shows a man and a car. Figure 5.16(a) is a perfect reference image. Figure 5.16(b) is an image created by blurring the car using a Gaussian blurring technique of radius 10. Figure 5.16(c) shows the same image but the blurring is done on the man. Figure 5.16 (d), (e) and (f) shows the fusion results of images on Figure 5.16(b) and (c) using the contourlet method, wavelet method and the simple averaging method.

Table 5.2 compares the quality of these composite images using WSSIM, PSNR and computational complexity between the original and the generated fused images. The result on Table 5.2 shows that the performance of the contourlet method is higher than that of the wavelet and simple averaging. The simple averaging technique takes the smallest amount of time as no selection procedure is done in the fusion process. The contourlet method takes the highest time.

Table 5.2 Performance evaluation result of the test image on Figure 5.16

Method	WSSIM	PSNR	Computational Time( in seconds)
<b>Contourlet Method</b>	<b>0.96344</b>	<b>30.0312</b>	<b>3.620078</b>
<b>Wavelet Method</b>	<b>0.96326</b>	<b>29.1992</b>	<b>3.392982</b>
<b>Simple Averaging</b>	<b>0.89271</b>	<b>24.6071</b>	<b>1.359040</b>

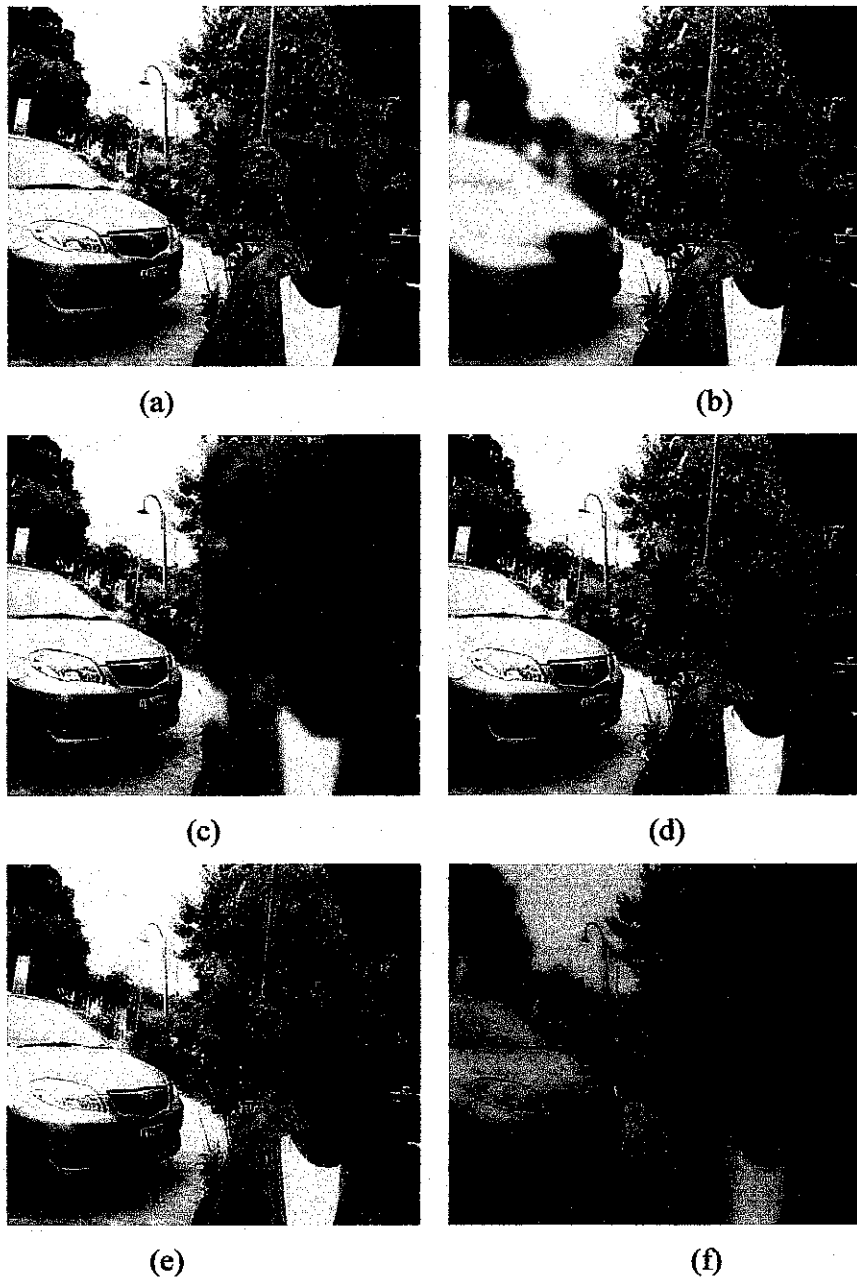


Figure 5.16 Objective evaluation test image (a) original reference image (b) right side blurred image (c) left side blurred image (d) contourlet method (e) wavelet method (f) simple averaging

## CHAPTER SIX: CONCLUSION

### 6.1 Introduction

This thesis developed innovative 8-bit color image enhancement techniques for improving the visibility of low quality digital images caused by high dynamic range, noise, poor contrast and low or non uniform illumination etc. The developed system can be used for security surveillance, driver's assistance at night or in low light condition, fire fighting operation, military and police target detection and acquisition, medical imaging, multimedia systems and search and rescue operations.

The underlying problems of image enhancement are solved using two approaches: a single image enhancement approach and composite image enhancement approach. The non linear approximation (NLA) systems using filter banks of wavelet and contourlet transform are studied thoroughly. Due to its efficiency in two dimensional data the contourlet transform is the major NLA method used in this work but the wavelet transforms are also implemented for comparison purposes. The objectives stated in chapter one are met:

1. The enhancement methods do not discard any salient information contained in the input image/images.
2. The methods do not introduce any artifacts or inconsistencies which can mislead a human observer or any subsequent image processing steps.
3. Both methods are reliable, robust and as much possible tolerant of imperfections such as noise.

The single image approach is highly dependent on the input image. If the input image is highly degraded and if the image details are covered by dark shadows, it fails to provide the needed enhancement. This problem is successfully solved by using the composite

image enhancement approach. This approach used two or more input images thus it increased the reliability by redundant information and capability by complementary information.

## **6.2 Summary of the Thesis**

This thesis gives some useful remarks about image processing and image representation in the computer system. A review of some of the most relevant literatures is also presented. Multiresolution representation approaches are briefly discussed. Two dimensional wavelet transform decomposition is first studied. The filter bank representation of the 2D wavelet decomposition is thoroughly presented. The contourlet transform is also discussed and formulated with the filter bank representation.

A novel single image enhancement algorithm is developed and proved to be efficient in several testing conditions. The technique for improving the visibility of low quality digital images caused by high dynamic range, noise, poor contrast and very low illumination is developed. The thesis also proposes a new approach on how to select a color space for optimal color image enhancement application. Ten most common color spaces are studied and evaluated.

At the end the thesis proposed a composite image enhancement approach for cases where the single image approach fails to provide the necessary enhancement. An image fusion system which is called fusion by selection of appropriate contourlet coefficients, which works well for both multi modal and multi focal images is developed. The fusion algorithm is successfully applied for combining several multi-modal and multi-focal images. The performance of the system is studied for different case study images.

## **LIST OF AWARDS**

1. **Melkamu Hunegnaw Asmare, Vijanth S. Asirvadam, Lila Iznita Izhar ,” CLEAR VU: Using Machine Vision, Image Enhancement and Fusion in Contourlet Domain”** SILVER Medal Award in Audio Visual Equipment Category, 20<sup>th</sup> International Invention Innovation and Technology Exhibition (ITEX09), Kuala Lumpur , Malaysia, 15-17 May 2009.
2. **Melkamu Hunegnaw Asmare , Vijanth S. Asirvadam, Lila Iznita Izhar,” Image Enhancement : Single and Composite Image Approaches using Contourlet Transform” ,** SILVER Medal Award in Postgraduate Category, 23<sup>rd</sup> Engineering Design Exhibition (EDX23), Universiti Teknologi PETRONAS, Malaysia, 22-23 April 2009.
3. **Melkamu Hunegnaw Asmare , Vijanth S. Asirvadam, Lila Iznitha, “ Multi-scale Color Image Enhancement using Contourlet Transform”,** Best Presenter Award, International Graduate Conference , IGCES2008, Johor Baharu , Malaysia, 23-24 Dec. 2008.
4. **Melkamu Hunegnaw Asmare , Vijanth S. Asirvadam, Lila Iznita Izhar,” Image Fusion For Improved Perception” ,** SILVER Medal Award in Postgraduate Category, 22<sup>nd</sup> Engineering Design Exhibition (EDX22), Universiti Teknologi PETRONAS, Malaysia, 15-16 October 2008.
5. **Melkamu Hunegnaw Asmare , Vijanth S. Asirvadam, Lila Iznita Izhar,” Image Fusion For Improved Perception” ,** BEST POSTER Award in Postgraduate Category, 22<sup>nd</sup> Engineering Design Exhibition (EDX22), Universiti Teknologi PETRONAS, Malaysia, 15-16 October 2008.



## LIST OF PUBLICATIONS

1. **Melkamu Hunegnaw Asmare, Vijanth S. Asirvadam, Lila Iznita Izhar** “Multi-scale Color Video Stream Enhancement Using Contourlet Transform”, International Conference on Science and Technolgiy application in Inductry and Education(ICSTIE2008), Penang,Malaysia, 12-13 Dec.08, pp : 675-678
2. **Melkamu Hunegnaw Asmare, Vijanth S. Asirvadam, Lila Iznita Izhar** “Multi-scale color image enhancement using contourlet transform”, International Graduate Conference in Engineering and Science (IGCES 08),Johor Bahru, Malaysia, 23-24 Dec.08, pp: 117-121.
3. **Melkamu Hunegnaw Asmare, Vijanth S. Asirvadam, Lila Iznita Izhar**, “Multi-Sensor Image Enhancement and Fusion for Vision Clarity Using Contourlet Transform”, International Confernce on Information Managemetn and Enginerring (ICIME09), April 3 - 5, 09, Kuala Lumpur, Malaysia, pp: 352-356.
4. **Melkamu Hunegnaw Asmare, Vijanth S. Asirvadam, Lila Iznita Izhar** ,“Color Space Selection for Color Image Enhancement Applications, International Conference on Signal Acquisition and Processing (ICSAP09), April 3 - 5,Kula Lumpur, Malaysia.pp:1-5.
5. **Melkamu Hunegnaw Asmare, Vijanth S. Asirvadam, Lila Iznita Izhar**, “Color Image Denoising and Contrast Enhancement in Contourlet Domain”, International Conference on Man-Machine Systems (ICoMMS 09), Kangar, Perlis, Malaysia, accepted, 11-13 October 2009.
6. **Melkamu H. Asmare, Vijanth S. Asirvadam, Lila Iznita Izhar**, “Image Enhancement : A Composite Image Approach”, International Conference on Electrical Engineering, (ICEEI09),accepted, August 5th – 7th, 2009 , Universiti Kebangsaan Malaysia, Malaysia.
7. **Melkamu H. Asmare, Vijanth S. Asirvadam, Lila Iznita Izhar**, “Multi-Scale Color Image Enhancement Using Contourlet Transform”, PLATFORM, July- December Issue, Universiti Teknologi PETRONAS, on Page(s): 145-151.

## BIBLIOGRAPHY

**A. F. Laine, J. Fan, S. Schuler**, "A framework for contrast enhancement by Dyadic Wavelet Analysis," Digital Mammography, A. G. Gale, S. M. Astley, D. R. Dance, and A. Y. Cairns, Eds. Amsterdam, The Netherlands: Elsevier, 1994, Page(s): 91-100.

**Andrea Masini, Francesco Branchitta, Marco Diani, Giovanni Corsini**, "Sight enhancement Through Video Fusion in a surveillance system" 14th International Conference on Image Analysis and Processing, 2007.

**Arthur L. Da Cunha, Jianping Zhou, Minh N. Do**, "The Nonsubsampled Contourlet Transform: Theory, Design, and Applications", IEEE Transaction on Image Processing, Volume 15, October 2006, Pages: 3089 -3101.

**D. Bhandari, Sankar K. Pal and Malay K. Kundu**, "Image Enhancement Incorporating Fuzzy Fitness Function in Genetic Algorithms" IEEE 1993, Pages: 1408-1413.

**D. Xiao, Jun Ohya**, "Contrast Enhancement of Color Images on Wavelet Transform and Human Visual system" Proceedings of IASTED international conference on Graphics and Visualization in Engineering, 2007, Page(s): 58-63.

**Dunkan Dun, Yin Po**, "Image Modeling in Contourlet Domain", MSc. Thesis, University of Illinois at Urbana – Champaign, 2003.

**E. J. Candes , D. L. Donoho** "Curvelets: a Surprisingly Effective Non-adaptive Representation for Objects with Edges" In A. Cohen, C. Rabut, and L. L. Schumaker, editors, Curve and Surface Fitting, Saint-Malo, 1999, Vanderbilt University Press.

**E. J. Candes , D. L. Donoho** "Curvelets, Multiresolution Representation, and Scaling Laws" In A. Aldroubi, A. F. Laine, and M. A. Unser, editors, SPIE Wavelet Applications in Signal and Image Processing VIII, Volume 4119, 2000.

**Edwin H. Land**, "Recent Advances in Retinex Theory and Some Implications for Cortical Computations: Color Vision and the Natural Image" Proc. National Academy Science USA, Vol. 80, August 1983, on Page(s): 5163-5169.

**E. Nezhadarya, Mohammad B. Shamsollahi and Omid Sayadi**, "Fuzzy Wavelet and Contourlet Based Contrast Enhancement" 28th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society, 2006. On Page(s): 6635-6638

**E. Nezhadarya, Mohammad B. Shamsollahi**, "Image Contrast Enhancement by Contourlet Transform" 48th International Symposium ELMAR-2006, 07-09 June 2006, Zadar, Croatia, Pages:81-84.

**Etta D. Piasano, Elodia B. Cole, Bradley M. Henminger**, "Image processing Algorithms for Digital Mammography: A pictorial Essay" Imaging and Therapeutic technology 2000, Volume 20, Page(s): 1479-1491.

**E. Hernandez, G. Weiss**, "A First Course on Wavelets", CRC Press, 1996.

**F. Sattar, Xinting Gao** "Image Enhancement Based on a Nonlinear Multiscale method using Dual-Tree Complex Wavelet Transform" IEEE Trans. Image Processing, June 1997, Volume 6, Page(s): 716-719.

**G. Apostolopoulos, E. Dermatas**, "Local Adaptive Contrast Enhancement in Digital Images", In Proceedings of the 3rd international Conference on Mobile Multimedia Communications (Nafpaktos, Greece, August 27 - 29, 2007). ACM International

Conference Proceeding Series, vol. 329, Institute for Computer Sciences Social-Informatics and Telecommunications Engineering, ICST, Brussels, Belgium, Page(s): 1-6.

**Gemma Piella Fenoy**, "Adaptive Wavelet and their Application in Image Fusion and Compression", PhD Thesis, Centre for Mathematics and Computer Science (CWI), Amsterdam, Oct. 2003.

**G. Petschnigg, M. Agrawala, H. Hoppe**, "Digital Photography with Flash and No-Flash Image Pairs", ACM 2004, Page(s): 664-672.

**G. Cheng, Li-Zhi Cheng**, "Adaptive Fingerprint Image Enhancement with Contourlet Transform" 2008 Congress on Image and Signal Processing, Pages: 261-264.

**H. Singh. Jyoti Raj. Gulsheen Kaur, Thomas Meitzler** "Image Fusion using Fuzzy Logic and Applications", IEEE International Conference Proceedings on Fuzzy Systems, 2004, Volume: 1, Page(s): 337- 340.

**H. Wang, Jun Wang, Wei Wang**, "Multispectral Image Fusion Approach based on GHM Multiwavelet Transform" Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21 August 2005, pages: 5043-5049.

**H. Wang**, "Multisensor Image Fusion by using Discrete Multiwavelet Transform" proceedings of the Third International Conference on Machine Learning and Cybernetics, Shanghai, 26-29 August 2004, pages: 4331-4336.

**Hui Xie, Weibin Rong, Lining Sun, Wei Chen**, "Image Fusion and 3-D Surface Reconstruction of Microparts using Complex Valued Wavelet Transforms" ICIP 2006, Pages: 2137-2140.

**Hummel, R. A.** "Image Enhancement by using histogram Transformation", Computer Vision, Graphics and Image Processing, 1977, Page(s): 174-184.

**Iyad Fayez Jafar**, "The Development and Evaluation of Novel Algorithms for Contrast Enhancement in Grayscale Images" PhD Thesis, Wayne State University, 2008.

**J. Qin and M. R. El-Sukku**, "A new Wavelet-based Method for Contrast/Edge Enhancement," in Proceedings of the 2003 International Conference on Image Processing, Volume 3, Barcelona, Spain, Sep.2003, Page(s): 397 – 400.

**Jean-Luc Starck, Fionn Murtagh, Emmanuel J. Candès, and David L. Donoho**, "Gray and Color Image Contrast Enhancement by the Curvelet Transform", IEEE Transactions on Image Processing, Volume 12, No 6, June 2003, on Page(s): 709-716.

**J. Tang, Qingling Sun, Kwaben Agyegon**, "An Image Enhancement Algorithm based on a Contrast Measure in the Wavelet Domain for Screening Mammograms" ICIP2007, Page(s): V29-V32.

**J. Singhai, P. Rawat**, "Image Enhancement method for underwater ,ground and satellite image using brightness preserving histogram equalization with maximum entropy" ICCIMA 2007, Page(s): 507-512.

**K. Barnard, B. Funt**, "Investigations into Multi-scale Retinex," in Color Imaging: Vision and Technology. New York: Wiley, 1999, on page(s): 9–17.

**Koen Vande Velde**, "Multi-scale Color Image Enhancement", IEEE Proceedings of International Conference on Image Processing (ICIP 99), 1999, Volume: 3, On Page(s): 584-587.

**K. Amolins, Yun Zhang, Peter Dare**, "Applications of Wavelet Transforms in Image Fusion" 2007 Urban Remote Sensing Joint Event, On page(s): 1-7

**K. Chen**, "Color Image Enhancement" ACM 30th Annual Southeast Conference, 1992, Page(s): 328-335.

**Li Tao**, "Multimodal Enhancement Technique for Visibility Improvement of Digital Images" PhD Thesis, Old Dominion University, 2005.

**Li Tao, Hau Ngo, Ming Zhang, Adam Livingston, Vijayan Asari**, "A Multi-sensor Image Fusion and Enhancement System for Assisting Drivers in Poor Lighting Conditions" Proceedings of the 34th Applied Imagery and Pattern Recognition Workshop (AIPR05), 2005, Page(s): 106-113.

**M. Hanmandlu, Devendra Jha, Rochak Sharma**, "Color image enhancement by fuzzy intensification" Pattern Recognition Letters volume 24, 2003, Pages: 81-87.

**M. N. Do** "Directional Multiresolution Image Representations" PhD thesis, Swiss Federal Institute of Technology, Lausanne, Switzerland, December 2001.

**M. N. Do, M. Vetterli, J. Stoeckler and G. V. Welland**, "Contourlets Beyond Wavelets" Academic Press, New York, 2002. Can be found in <http://www.ifp.uiuc.edu/~minhdo/publications>.

**M. N. Do, M. Vetterli** "Pyramidal directional Filter banks and Curvelets" In Proceeding of IEEE International Conference on Image Processing, Thessaloniki, Greece, Oct. 2001.

**M. N. Do, M Vetterli**, "The Contourlet Transform: An efficient Directional Multi Resolution Image Representation", IEEE Transactions on Image Processing, 2005, Volume 14, Pages: 2091-2106.

- M. Vetterli**, "Wavelets, Approximation, and Compression", September 2001 IEEE Signal Processing Magazine, Pages: 59 -73.
- Minh N. Do**, "Contourlets and Sparse Image Expansions", SPIE 2003, Volume 5207, Page(s): 560-570.
- Monica A. Trifans, John M. Tyler, Oleg S. Pianyky**, " Applying Multiresolution method to medical image Enhancement" ACM SE 2006, Page(s): 254-259.
- N. Humphrey**, "Cave Art, Autism, and the Evolution of Human Mind", Cambridge Archaeological Journal 8:2, 1998, Page(s): 164-191.
- OCTEC Limited**, "Applied Image Processing Technololgy, Design Manufacturing and Image Processing Sub-systems", can be reached on online: <http://www.octec.co.uk/>.
- Oliver Rockinger**, "Metapix, Solutions Beyond the Pictures" available online on : <http://www.metapix.de>
- P. J. Burt and E. H. Adelson**, "The Laplacian pyramid as a Compact Image Code" IEEE Transaction in Communication , April 1983, volume 31, Pages:532-540.
- P. J. Burt**, "The pyramid as a structure for efficient computation" In Multiresolution Image Processing and Analysis, A. Rosenfeld, Ed. Springer-Verlag, Berlin, 1984, Page(s): 6-35.
- P. J. Burt, Koleczynski, R. J.** "Enhanced image capture through fusion" In Proceedings of the 4th International Conference on Computer Vision (Berlin, Germany, May 1993, Pages: 173- 182.

**P. Dong-Liang, Xue Anke**, "Degraded Image Enhancement with Applications in Robot Vision" IEEE conference on systems, Man and cybernetics, 2005, Volume 2, Page(s): 1837-1842.

**P. Feng , Yingjun Pan, Biao Wei, Wei Jin, Deling Mi** , " Enhancing Retinal Images by the contourlet Transform" Pattern Recognition Letters, Volume 28 , 2007 , Page(s): 512-522.

**Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., Romeny, B. H., Zimmerman, J. B., Karel, Z.**, "Adaptive histogram equalization and its variation," Computer Vision, Graphics and Image Processing, Volume 39, 1987, on Page(s): 355-368.

**R. A. DeVore, B. Jawerth, B. J. Lucier** "Image Compression through Wavelet transform coding" IEEE Transactions on Information Theory, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis, March 1992, Volume 38 Pages:719-746.

**R. C. Gonzalez, R. E. Woods**, "Digital Image Processing." Addison Wesley, 1992. (Chapter 4)

**R. H. Bamberger, M. J. T. Smith** "A Filter bank for the directional decomposition of images: Theory and design" IEEE Transactions in Signal Processing, April 1992, volume 40, Pages: 882-893.

**Ranchin, T., Wald, L.** "Fusion of High Spatial and Spectral Resolution Images: the ARSIS concept and its implementation" Photogrammetric Engineering and Remote Sensing, Volume 66, 2000, Page(s): 49-61.

**Ronald A. Devore**, "Nonlinear Approximation", Acta Numerica (1998), Cambridge University Press, Page(s): 51-150.



**R. Yang, Quan Pan, Yong-mei Cheng**, "A new method of Image Fusion based on m-band Wavelet Transform" Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 13-16 August 2006, Pages: 3870-3873.

**S. G. Nikolov, D. R. Bull, C. N. Canagarajah, M. Halliwell, P. N. T. Wells**, "Image Fusion using 3-D Wavelet Transform" Image Processing and its Applications, Conference Publication No. 465, IEE 1999, Pages: 235-239

**S. Mallat**, "A theory for Multiresolution Signal Decomposition: the Wavelet Representation", IEEE Trans. Pattern Recognition and Machine Intelligence, July 1989, Volume 11, Page(s): 674-693.

**S. Mallat**, "A wavelet Tour of Signal Processing", Academic Press, 2<sup>nd</sup> Edition, 1999.

**S. T. Li, Y.N. Wang**, " Multisensor Image Fusion using Discrete Multiwavelet Transform" Proceedings of the 3<sup>rd</sup> International Conference on Visual Computing, Mexico City , Mexico, 2000, Volume: 7, On page(s): 4331- 4336.

**Shojiro Maki, Toshimitsu Konno, Hiroshi Maeda**, "Image Enhancement in Computerized Tomography for Sensitive Diagnosis of Liver Cancer and Semi quantization of Tumor Selective Drug Targeting With Oily Contrast Medium" Cancer volume 56, issue 4, 2006, Pages :751-757.

**Toet, A.** "Image fusion by a ratio of low-pass pyramid" Pattern Recognition Volume 9, 1989, Pages: 245-253.

**W.J. Phillips**, "Wavelets and Filter Banks" Course Notes, January 9, 2003, can be found online on: <http://www.engmath.dal.ca/courses/engm6610/notes/notes.html>.

**Xuli Zong, A.F. Laine, E. A. Geiser, D.C. Wilson**, “De-Noising and contrast enhancement via wavelet shrinkage and nonlinear adaptive gain” Proceedings of SPIE, in Wavelet Applications III, 1996. Volume 27, on Page(s): 566-574.

**Y. Meyer, R. D. Ryan** “Wavelets: Algorithms and Applications” Society for Industrial and Applied Mathematics, Philadelphia, 1992.

**Yi Chen, Michael D. Adams, and Wu-Sheng Lu**, “Design of Optimal Quincunx Filter Banks for Image Coding” ISCAS 2006, Page(s): 2041-2044.

**Yi Yang, Chong-zhao Han, De-qiang Han**, “A Structure information based Image Fusion Algorithm using IHS and Discrete Wavelet Transform” Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China, 2-4 Nov. 2007, Pages : 859-864.

**Yin Chen**, “Investigation of Image Fusion Algorithms for Performance Evaluation for Night Vision”, PhD Thesis, Lehigh University, March 2007.

**Yinpeng Jina, Laura Fayad, Andrew Laine**, “Contrast Enhancement by Multi-scale Adaptive Histogram Equalization” Proceedings of SPIE, Volume 4478, 2001, Page(s):206-213.

**Z. Rahman, D. J. Jobson, G. A. Woodell**, “Multi-scale Retinex for Color Image Enhancement,” IEEE International Conference on Image Processing, 1996. Vol. 3, on Page(s): 1003-1006.

**Zhengyou Wang, Quan Xue, Guobin Chen, Weiming Zeng, Zhijun Fang, Shiqian Wu** , “A New Color Image Enhancement Algorithm for Camera-Equipped Mobile Telephone” KES 2006, Part I, LNAI 42512006, Page(s): 631 – 638.

**Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, Eero P. Simoncelli**, "Image Quality Assessment: From Error Visibility to Structural Similarity", IEEE Transaction on Image Processing, Volume 13, NO. 4, April 2004, Page(s): 600-612.

## APPENDIX A: DERIVATION OF THE ENHANCEMENT FUNCTION

Let  $p(t) = \frac{1}{1 + ce^{-t^2}}$  where  $c$  is a constant which controls the shape of the function.  $p(t)$  is a member of the logistic function of the sigmoid curve. The sigmoid curve has a characteristic shape of the letter 'S'. But for enhancement purposes we need the inverted 'S' shape thus we will use the reciprocal of function  $p(t)$ .

Let  $y(t) = \tanh(bt) = \frac{e^{2bt} - 1}{e^{2bt} + 1}$  where  $b$  is a constant which controls the shape of the function. The hyperbolic tangent function is also a member of the sigmoid curve.

Then the enhancement function  $E(t)$  can be derived by combining the two functions as follows:

$$E(t) = f(y(t), p(t))$$

One possible way to combine these two functions is by defining parameters  $\alpha > 0$  and  $\beta > 0$ , which adjust the importance of the component functions. Thus

$$E(t) = [y(t)]^\alpha \cdot [p(t)]^{-\beta}$$

For simplicity let  $\alpha = \beta = 1$ , then:  $E(t) = \frac{e^{2bt} - 1}{e^{2bt} + 1} \times \left( \frac{1}{1 + ce^{-t^2}} \right)^{-1}$ ,

$$E(t) = \frac{(e^{2bt} + 1)(1 + ce^{-t^2})}{e^{2bt} + 1}$$

Simplifying the above function we get:

$$E(t) = \frac{e^{2bt} - ce^{-t^2} + ce^{2bt-t^2} - 1}{e^{2bt} + 1}$$

Which is the enhancement function used in chapter four.

APPENDIX B: THE GRAPHICAL USER INTERFACE (GUI)

B.1 GUI for Single Image Enhancement Approach

The GUI for single image enhancement approach is shown below. It is furnished with several functions all of which are written using MATLAB 7.1. The user can load an image. Four different enhancement modes can be selected. The gain and denoising drop down buttons work for wavelet and contourlet methods and control the enhancement function.

The user can compare different enhancement results by using the histogram equalization and profile of the image. It is possible to enlarge the image for a better display and save the enhanced image.

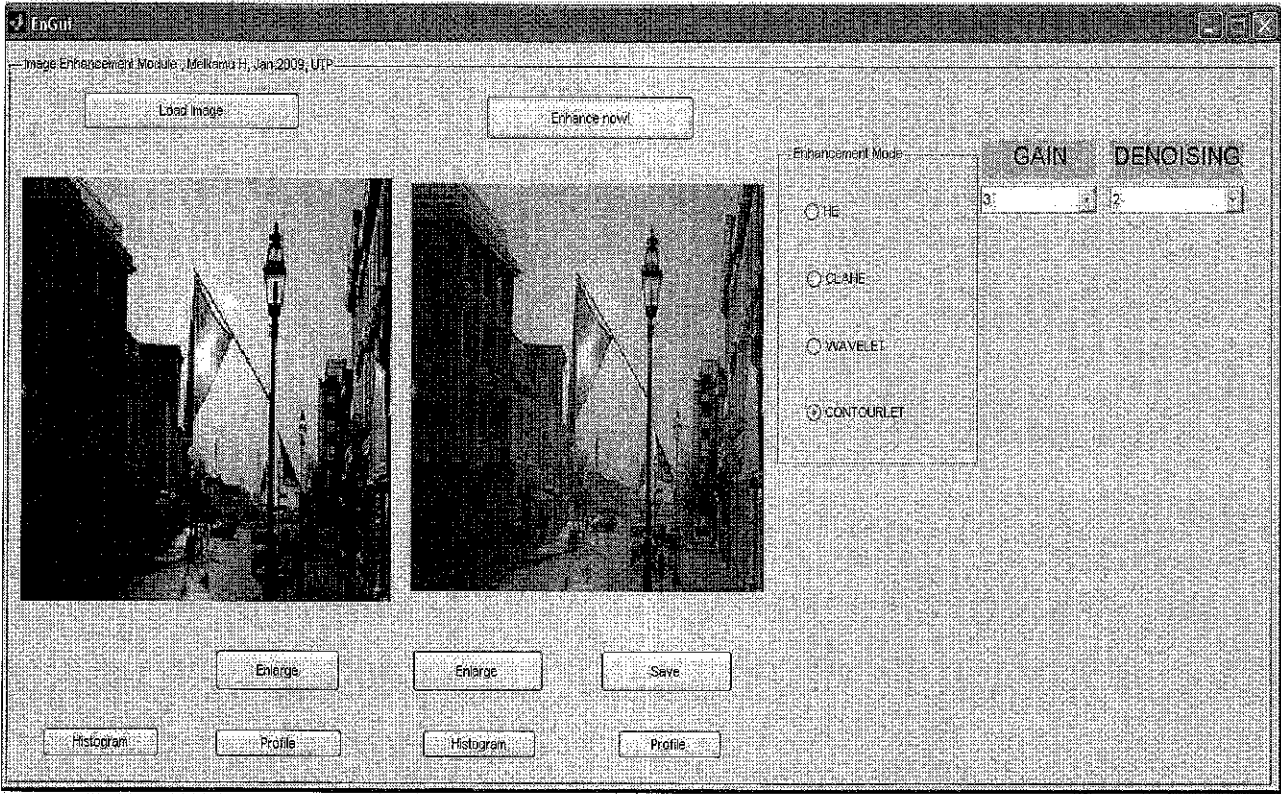


Figure B.1 GUI for single image

B.2 GUI for Composite Image Approach

This GUI is also a collection of several functions written using MATLAB 7.1. In this case the user loads two images of the same size and pre aligned. Two Multiresolution analysis methods (wavelet and contourlet) can be chosen. The level of decomposition is also selected by the user. The combination tab tells the user to select between pixel based and region based fusion approaches. However, in this version only the pixel based method is implemented.

The user can see the enlarged view by using the view button and can save the enhanced image.

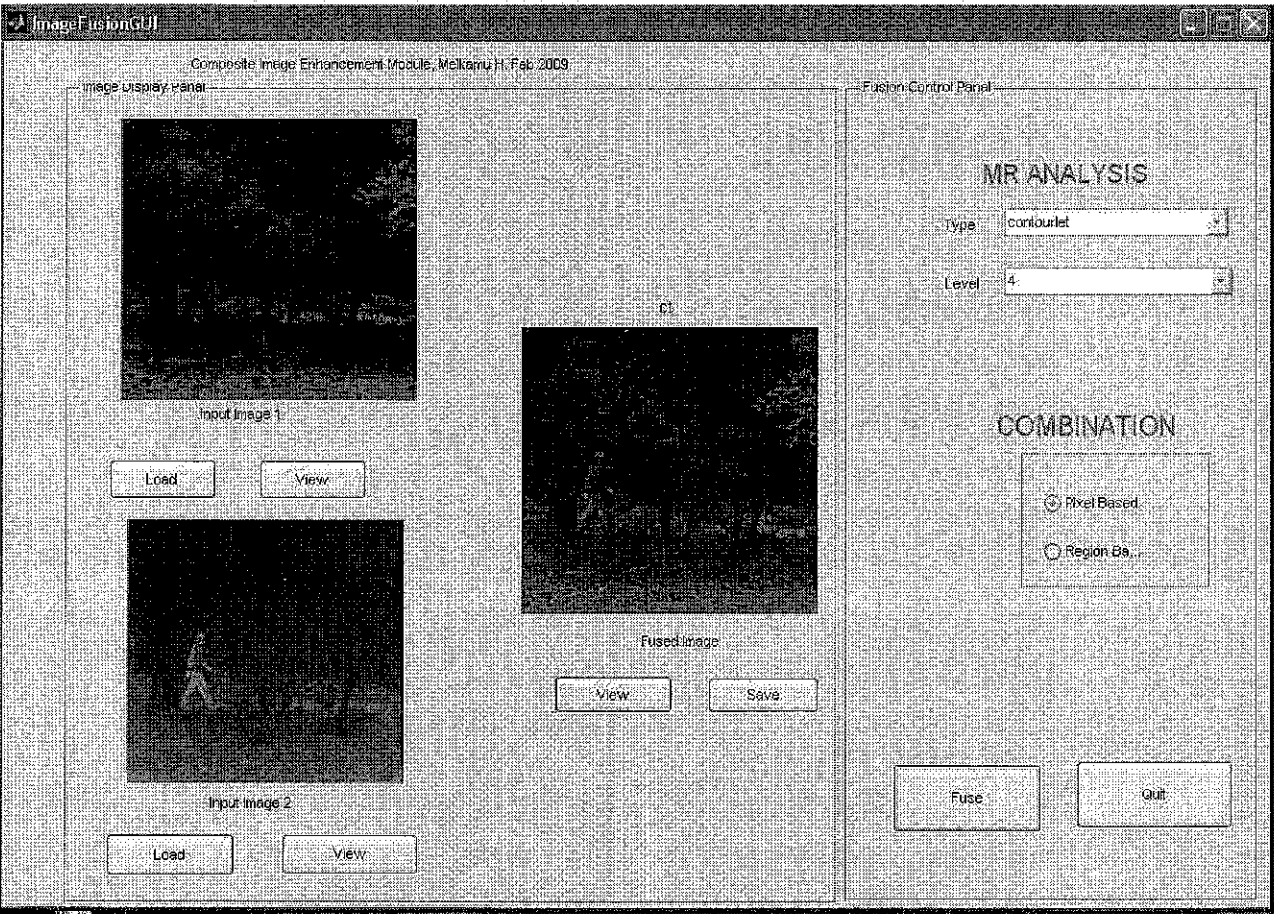


Figure B.2 GUI for composite image