Multifocus Image Fusion based on lucidity Decision Parameters

By

Nasrulla Nazir

[2006-NUST-MS PhD-CSE (E)-25]

Submitted to the Department of Computer Engineering
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Software Engineering

Advisor
Brig. Dr Muhammad Younus Javed

College of Electrical & Mechanical Engineering
National University of Sciences and Technology

2009

In the name of ALLAH, the most Beneficent, the most Merciful.

ABSTRACT

The goal of image fusion is to combine information from multiple images of the same scene. The result of image fusion is a single image which is more suitable for human perception or further image processing tasks. It is obtained by extracting all the useful information from the source images while not introducing artifacts or inconsistencies which can distract human observers or the following processing. Nowadays, image fusion has become an emerging and essential tool and shown its power in many fields like image analysis and computer vision, automatic object detection, robotics, military and law enforcement, satellite imagery, night vision applications, remote sensing and medical diagnosis. For this purpose a novel image fusion technique has been proposed. Firstly original input multi-focus images are partitioned into blocks. Then clarity of these blocks is decided on the basis of three distinctive features i.e. Spatial Frequency, Image Clarity and Block Visibility. Using this decision, the original input multi-focus images are further decomposed into much smaller blocks and again decision is made for blocks which are on the boundary of focused and blurred portions on the basis of the same three distinguishing features. After this practice, all focused and blurred blocks of the original images are clearly identified. Then the smaller blocks which are on the boundary of the clear and blurred parts are fused using conventional wavelet transform and all other blocks away from boundary are taken from original images as intact.

Experimental results on standard test images (i.e. Lena, Barbara and Peppers) clearly depict that the proposed approach outshines classical discrete wavelet transform based image fusion techniques and many other. These results provide higher Peak Signal to Noise Ratio (PSNR) and smaller Root Mean Square Error (RMSE) values than some of the previous approaches.

MATLAB 7.0 has been used for the implementation of the proposed approach. Experiments have been carried out on a variety of standard grayscale images with different defocus parts.

Dedications

To all those who believe in ….

“No one gets more than what he strives for”

(Al Quran)

ACKNOWLEDGEMENTS

All praise be to the omnipotent and magnanimous ALMIGHTY ALLAH, who guides us from darkness to the light and help us in difficulties, who enable me to complete this tough work and bestowed the wisdom and strength to make material contribution to an already existing ocean of knowledge. I offer our humblest thanks to the prophet Muhammad (peace be upon him), who is forever a source of guidance and knowledge for humanity.

I do not wish to merely complete a formality when expressing my gratitude and admiration for my supervisor Brig. Dr. Muhammad Younus Javed. Apart from his skilled supervision and invaluable help in the project work, I shall always be indebted to him for the helping hand he lend me in the hours of crisis. At the same time I equally express my respect and regards to the all the committee members for their time and diligence with whom I had quite a few good discussions and their never ending queries sharpened my own concepts.

I express special thanks to my friends, especially Mr. Muhammad Iqbal, for providing unconditional and continuous support during time of crisis.

Finally, I want to thank my beloved parents and my wife for their continuous encouragement, understanding and moral support; and have a deep gratitude to them.

Table of Contents

ABSTRACT III
ACKNOWLEDGEMENTS V
TABLE OF CONTENTS VI
INTRODUCTION 1
1.1 IMAGE FUSION 1
1.2 GENERALIZATION IN THE FIELD OF IMAGE FUSION 2
1.2.1 Pixel Level Image Fusion 2
1.2.2 Region Level Image Fusion 3
1.2.3 Decision Level Image Fusion 4
1.3 IMAGE FUSION METHODS 4
1.3.1 Wavelet Transform 4
1.3.2 Multiresolution or Multi-scale Methods 5
1.3.3 Image Pyramids 5
1.3.4 Artificial Neural Networks 6
1.3.5 Linear Superposition 6
1.3.6 Non-Linear Methods 6
1.3.7 Generic Multiresolution Scheme 7
1.4 MATLAB 8
1.5 SCOPE OF THESIS 9
1.6 THESIS OUTLINE 9
1.7 SUMMARY 10
MULTIFOCUS IMAGE FUSION 11
2.1 MULTIFOCUS IMAGES 11
2.2 MULTIFOCUS IMAGE FUSION 11
2.3 MULTIFOCUS IMAGE FUSION METHODS 12
2.3.1 Region Selection Methods 13
2.3.2 Learning Based Methods 13
2.3.3 Multi-scale Decomposition (MSD) Methods 13
2.4 ADVANTAGES OF MULTIFOCUS IMAGE FUSION 14
2.5 RELATED RESEARCH FIELDS 14
2.5.1 Intelligent Robots 14
2.5.2 Medical Imaging 15
2.5.3 Manufacturing 15
2.5.4 Military and Law Enforcement 15
2.5.5 Remote Sensing 15
2.6 SUMMARY 15
FEATURE LEVEL IMAGE FUSION 16
3.1 OVERVIEW OF FEATURE LEVEL IMAGE FUSION 16
3.2 DECISION FEATURES EXTRACTIONS METHODS 17
3.3 TRANSFORMS 17

3.3.1 Overview of Wavelet Transform *18*
3.3.2 What is a Wavelet? *18*
3.3.2.1 Scaling 19
3.3.2.2 Shifting 20
3.3.2.3 Scale and Frequency 20
3.3.3 One-dimensional Discrete Wavelet Transform *21*
3.3.4 Two-dimensional Discrete Wavelet Transform *23*

3.4 ADVANTAGES OF FEATURE LEVEL IMAGE FUSION 25

3.5 SUMMARY 25

DECISION FEATURES 26

4.1 OVERVIEW 26

4.2 DECISION PARAMETERS USED IN PROPOSED APPROACH 27

4.2.1 Image Clarity *27*
4.2.2 Spatial Frequency *27*
4.2.3 Block Visibility *28*

4.3 SUMMARY 28

IMPLEMENTATION METHODOLOGY 29

5.1 DESIGN OF SYSTEM 29

5.1.1 Description of proposed scheme *30*
Step 1- Image Acquisition 31
Step 2- Application of Average Filter 31
Step 3- Partitioning both multi-focused images into blocks of order 16x16 33
Step 4- Measuring Clarity 33
Step 5- Applying Majority Filter on DM_{16} 34
Step 6- Is Boundary Block? 35
Step 7- Applying Majority Filter on DM8 36
Step 8- Is Boundary Block? 37
Step 9 – Reconstruction of fused image 37

5.2 THE SIMULATION 38

5.3 SUMMARY 45

RESULTS AND DISCUSSION 46

6.1 PERFORMANCE MEASURES 46

6.1.1 Peak Signal to Noise Ratio (PSNR) *46*
6.1.2 Root Mean Square Error (RMSE) *47*

6.2 EXPERIMENTAL RESULTS 47

6.3 SUMMARY 51

CONCLUSION AND FUTURE WORK 52

7.1 CONCLUSION 52

7.2 FUTURE WORK 53

REFERENCES 54

LIST OF FIGURES

FIGURE 1.1: THE GENERIC MULTIRESOLUTION SCHEME 8
FIGURE 2.1: MULTIFOCUS IMAGE FUSION OF LENA IMAGE 12
FIGURE 3.1: COMPARISON OF A SINE WAVE AND A WAVELET [34] 19
FIGURE 3.2: EFFECT OF SCALE FACTOR ON A WAVELET [34] 20
FIGURE 3.3: EFFECT OF SHIFTING A WAVELET [34] 20
FIGURE 3.4: RELATIONSHIP BETWEEN SCALE AND FREQUENCY OF A WAVELET [34] 21
FIGURE 3.5: AN FWT ANALYSIS BANK 22
FIGURE 3.6: MULTILEVEL WAVELET-DECOMPOSITION OF A SIGNAL 23
FIGURE 3.7: THE RESULTING DECOMPOSITION OF A TWO-DIMENSIONAL FWT [35] 24
FIGURE 5.1 FLOW CHART OF PROPOSED SCHEME 29
FIGURE 5.2: ORIGINAL LENA TEST IMAGE OF SIZE 512X512 31
FIGURE 5.3: MULTI-FOCUSED IMAGES GENERATED USING AVERAGE FILTER FROM LENA TEST IMAGE 33
FIGURE 5.4: DECISION MAP (DM_{16}) OF UPPER AND LOWER MULTI-FOCUSED LENA IMAGE 34
FIGURE 5.5: DECISION MAP (DM_{16}) OF UPPER AND LOWER MULTI-FOCUSED LENA IMAGE 35
AFTER APPLYING MAJORITY FILTER 35
FIGURE 5.6: DECISION MAP (DM_8) OF UPPER AND LOWER MULTI-FOCUSED LENA IMAGE 36
FIGURE 5.7: DECISION MAP (DM_8) OF UPPER AND LOWER MULTI-FOCUSED LENA IMAGE 36
AFTER APPLYING MAJORITY FILTER 36
FIGURE 5.8: THE RESULTING FUSED IMAGE 38
FIGURE 5.9: SCREEN SHOT OF DATA ACQUISITION 39
FIGURE 5.10: SCREEN SHOT OF MULTI-FOCUSING OPTIONS 40
FIGURE 5.11: SCREEN SHOT OF DECISION MAP 16X16 41
FIGURE 5.12: SCREEN SHOT OF DECISION MAP 16X16 AFTER APPLYING MAJORITY FILTER 42
FIGURE 5.13: SCREEN SHOT OF DECISION MAP 8X8 43
FIGURE 5.14: SCREEN SHOT OF DECISION 8X8 AFTER APPLYING MAJORITY FILTER 44
FIGURE 5.15: SCREEN SHOT OF RECONSTRUCTED IMAGE 45
FIGURE 6.1: COMPARISON OF RMSE 50
FIGURE 6.2: COMPARISON OF PSNR 50

LIST OF TABLES

TABLE 6.1: COMPARISON OF FUSED IMAGE WIT H INPUT IMAGES OF LENA.. 47
TABLE 6.2: COMPARISON OF FUSED IMAGE WIT H INPUT IMAGES OF PEPPERS.. 48
TABLE 6.3: COMPARISON OF FUSED IMAGE WIT H INPUT IMAGES OF BARBARA 48
TABLE 6.4: IMAGE FUSION METHODS' COMPARISONS FOR LENA IMAGE.. 49
TABLE 6.5: IMAGE FUSION METHODS' COMPARISONS FOR BARBARA IMAGE... 49
TABLE 6.6: IMAGE FUSION METHODS' COMPARISONS FOR PEPPERS IMAGE 49

Chapter 1

Introduction

1.1 Image Fusion

Image fusion is a technology that combines the information of the same scene of two or more pictures from identical or different types of sensors to generate a more precise, comprehensive and reliable image description or interpretation of that scene (usually a composite image), which is more suitable for the purposes of human visual perception or computer-processing tasks such as segmentation, feature extraction and target recognition [1]. Multi-source image fusion, different from general image enhancement, is a new technology in the fields of computer vision and image comprehension. Image fusion fully utilizes much complementary and redundant information of the original images. 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. Generally there are three requirements of the image fusion algorithm:

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

The benefits of image fusion include:

i. Extended range of operations: multiple sensors that operate under different operating conditions can be deployed to extend the effective range of operations. For example different sensors can be used for day/night operation.

ii. Extended spatial and temporal coverage: joint information from sensors that differ in spatial resolution can increase the spatial coverage. The same is true for the temporal dimension.
iii. Reduced uncertainty: joint information from multiple sensors can reduce the uncertainty associated with the sensing or decision process.
iv. Increased reliability: the fusion of multiple measurements can reduce noise and therefore improve the reliability of the measured quantity.
v. Robust system performance: redundancy in multiple measurements can help in the system's robustness. In case one or more sensors fail or the performance of a particular sensor deteriorates, the system can depend on the other sensors.
vi. Compact representation of information: fusion leads to compact representations. For example, in remote sensing, instead of storing imagery from several spectral bands, it is comparatively more efficient to store the fused information.

1.2 Generalization in the field of Image Fusion

There are three main abstraction levels in Image fusion processes which are pixel, region and decision based levels depending on the processing phenomena for image fusion. Pixel level image fusion operates on several sources of raw data (pixels) to produce new raw data that is expected to be more informative and synthetic than the inputs. Region level fusion methods divide an image in segments based on some criteria such as edges, corners, lines, texture parameters, etc and than based on this information, a feature map is built which is used for further processing. Decision level fusion takes results from several pre defined algorithms, process hem and produce final better decision about fused image.

1.2.1 Pixel Level Image Fusion

Typically, in image processing, images presenting several spectral bands of the same scene are fused to produce a new image that ideally contains in a single channel all/(most) of the information available in the various spectral bands. An operator (or an image processing algorithm) could then use this single image instead of the original images. This is particularly important when the number of available spectral bands becomes so large that it is impossible to look at the images separately. This kind of fusion requires a precise (pixel-level) registration of the available images. This registration is intrinsic when the various bands come from the same sensor but it is a lot more complicated when several different sensors are used (SAR, IR scanner, camera, etc.). Pixel level fusion combines several sources of raw data to produce new raw data that is expected to be more informative and synthetic than the inputs [12-14].

1.2.2 Region Level Image Fusion

Region level image fusion generally involves that an image is initially segmented in some way to produce a set of regions and then various properties of these regions can be calculated and used to determine which features from which images are included in the fused image. In most applications of image fusion, people pay more attention to fused objects rather than individual pixels. Since objects could be represented by regions, thus region-based fusion approaches could be more meaningful than pixel-based fusion methods. Meanwhile region-based fusion scheme has the advantages of reduced sensitivity to noise, blurring effects and misregistration. A number of region-based fusion schemes have been proposed. These methods initially transform the pre-registered images to multi-scale representations, and regions are extracted from each source image. The source images are then fused based on a simple region property such as average activity. These methods do not take full advantage of the wealth of information that can be calculated for each region.

Region level image fusion also named as Feature level image fusion combines various features. These features may come from several raw data sources (several sensors, different moments, etc.) or from the same raw data. In the latter case, the objective is to find relevant features among available features that might come from several feature extraction methods. The objective is to obtain a limited number of relevant features. Typically, in image processing, feature maps are computed as pre-processing for segmentation or detection. Features such as edges, corners, lines, texture parameters

(Haralick, Wavelet coefficients, etc.) are computed and combined in a fused feature map that may then be used for segmentation or detection.

1.2.3 Decision Level Image Fusion

Decision level fusion combines decisions coming from several algorithms and experts. Methods of decision levels fusion include voting methods, statistical methods, fuzzy logic based methods, etc.

The above categorization does not encompass all possible fusion paradigms, as input and output of the fusion process may present different levels of processing. Typically features could be fused to output a decision. In practical problems, the applied fusion procedure is often a combination of the previously mentioned three levels.

1.3 Image Fusion Methods

In recent years, many research achievements are gotten in image fusion field. The most important issue concerning image fusion is to determine how to combine the input images. In recent years, several image fusion techniques have been proposed. The primitive fusion schemes perform the fusion right on the source images. One of the simplest of these image fusion methods just takes the pixel-by-pixel gray level average of the source images. This simplistic approach often has serious side effects such as reducing the contrast. Some of the prominent image fusion methods are discussed in the following subsections.

1.3.1 Wavelet Transform

An image analysis method similar to image pyramid is the discrete wavelet transform [17-20]. Wavelet transform is a new method for multiresolution analysis, by which an image can be decomposed into the lowest approximation and several details at different scales and in different directions. The lowest approximation contains the average information and most energy of the image, while details contain edges or high frequency information at different scales and in different directions. Among the multifocus images, there are certain objects clear in some images, and some other objects clear in other images. The magnitudes of the wavelet coefficients for blurred area are less than that for clear area. So for the high frequency Coefficients (wavelet coefficients), we simply take the coefficient that is of the greatest magnitude among the multifocus images as that of the fusion image, while for the lowest approximation, we take the average as usual.

Apart from its various advantages, two main drawbacks of the DWT are the existence of shift variance and the directional constraint in diagonal feature extraction (45◦ plane). One of the oldest transforms compensating the DWT disadvantages is the shift-invariant DWT (SIDWT), which is based on the fact that not all shifts are necessary for perfect signal.

1.3.2 Multiresolution or Multi-scale Methods

In these methods, the input images are decomposed into their multi-scale edge representation, using either any image pyramid or any wavelet transform. The actual fusion process takes place in the synthesis, where the fused multi-scale representation is built by a pixel-by-pixel selection of the coefficients with maximum magnitude. Finally the fused image is computed by an application of the appropriate reconstruction scheme. These techniques represent unique information at different resolutions and are able to reconstruct signals perfectly. Image fusion based on these techniques extracts important features of every resolution and improves the image quality effectively.

1.3.3 Image Pyramids

A generic image pyramid is a sequence of images where each image is constructed by low pass filtering and sub sampling from its predecessor [15-16]. The most popular methods are based on the pyramid transform (such as Laplacian pyramid, ratio pyramid, gradient pyramid), the wavelet transform, the contourlet transform and so on. Image fusion based on the pyramid transform decomposes a image from coarse to fine, and at each level generates a band pass image. The process is iterated on the coarse signal. This kind of fusion method can obtain good visual results. But pyramid transform has redundancy, the image data after decomposition is increased and the fused image is not the best. The wavelet transform is widely used in image fusion field and many research achievements are gotten. Wavelets have shown their ability in representing natural images that contain smooth areas separated with edges. However, natural images consist of edges that are smooth curves and which cannot be captured efficiently by the wavelet transform.

1.3.4 Artificial Neural Networks

Inspired by the fusion of different sensor signals in biological systems, many researchers have employed artificial neural networks in the process of image fusion [2-5]. The most popular example for the fusion of different imaging sensors in biological systems is described by Newman and Hartline in the 80s: Rattlesnakes (and the general family of pit vipers) possess so called pit organs which are sensitive to thermal radiation through a dense network of nerve fibers. The output of these pit organs is fed to the optical tectum, where it is combined with the nerve signals obtained from the eyes. Newman and Hartline distinguished six different types of bimodal neurons merging the two signals based on a sophisticated combination of suppression and enhancement.

Several researchers modeled this fusion process for the combination of multi spectral imagery by a combination of several neural networks.

1.3.5 Linear Superposition

Probably, the most straightforward way to build a fused image of several input images is performing the fusion as a weighted superposition of all input images.

The optimal weighting coefficients, with respect to information content and redundancy removal, can be determined by a principal component analysis (PCA) of all input intensities. By performing a PCA of the covariance matrix of input intensities, the weightings for each input frame are obtained from the eigenvector corresponding to the largest Eigen value.

1.3.6 Non-Linear Methods

Another simple approach to image fusion is to build the fused image by the application of a simple nonlinear operator such as max or min. If in all input images the bright objects are of interest, a good choice is to compute the fused image by a pixel-by-pixel application of the maximum operator.

An extension to this approach follows by the introduction of morphological operators such as opening or closing. One application is the use of conditional morphological operators by the definition of highly reliable 'core' features present in both images and a set of 'potential' features present only in one source, where the actual fusion process is performed by the application of conditional erosion and dilation operators.

A further extension to this approach is image algebra, which is a high-level algebraic extension of image morphology, designed to describe all image processing operations. The basic types defined in image algebra are value sets, coordinate sets which allow the integration of different resolutions and tessellations, images and templates. For each basic type, binary and unary operations are defined which reach from the basic set operations to more complex ones for the operations on images and templates. Image algebra has been used in a generic way to combine multi-sensor images.

1.3.7 Generic Multiresolution Scheme

The basic idea of the generic multiresolution fusion scheme is motivated by the fact that the human visual system is generally sensitive to local contrast changes, i.e. edges. Motivated from this insight, and in mind that both image pyramids and the wavelet transform result in a multiresolution edge representation, it is straightforward to build the fused image as a fused multiscale edge representation.

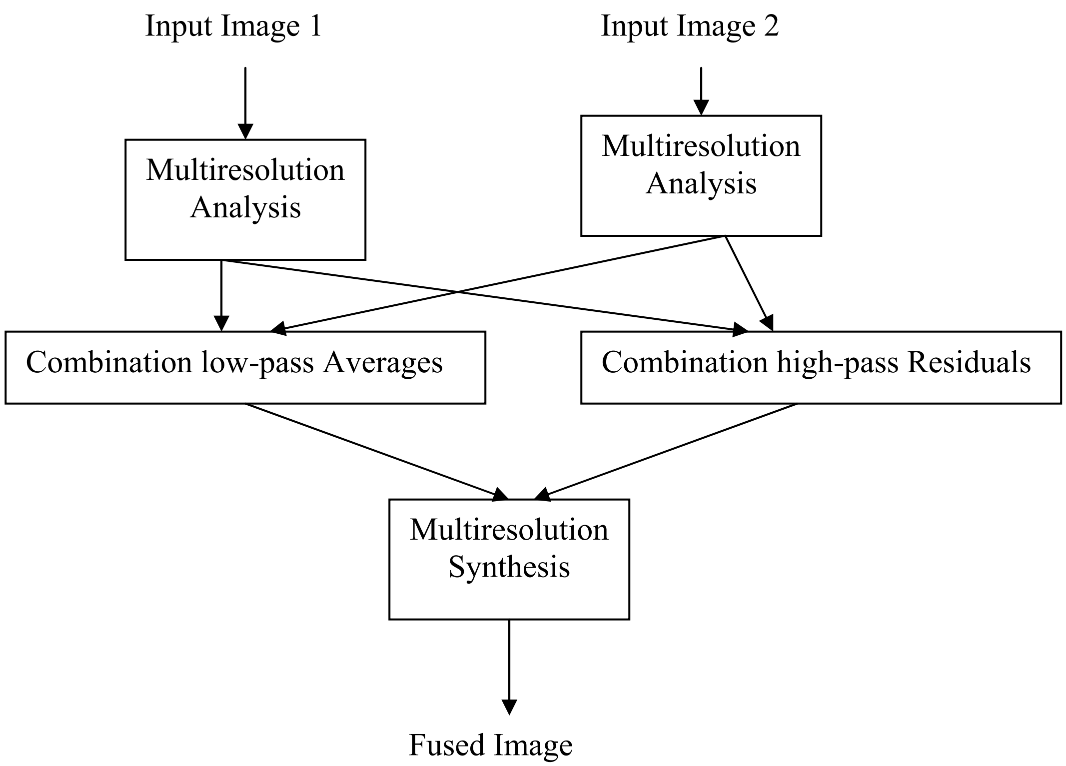


Figure 1.1: The generic multiresolution scheme

The fusion process is summarized in the Figure 1.1: In the first step the input images are decomposed into their multiscale edge representation, using either any image pyramid or any wavelet transform. The actual fusion process takes place in the synthesis, where the fused multiscale representation is built by a pixel-by-pixel selection of the coefficients with maximum magnitude. Finally the fused image is computed by an application of the appropriate reconstruction scheme.

1.4 MATLAB

This research has been completely implemented in MATLAB®. This high performance language for technical computer, integrates computation, visualization, and programming in an easy-to-use environment. One of the reasons of selecting MATLAB® in this research is that its original concept of a small and handy tool which has evolved to become an engineering workhorse. It is now accepted that MATLAB and its numerous Toolboxes can replace and/or enhance the usage of traditional simulation tools for advanced engineering applications. And it also fits perfectly in the necessities of an image processing research due to its inherent characteristics. Basic data element of MATLAB® is an array that does not require dimensioning. This is especially helpful to solve problems with matrix and vector formulations. And an image is nothing but a matrix or a set of matrices which defines the pixels value of the image, such as grey scale value in black and white images, and Red, Green and Blue or Hue, Saturation and Intensity values in color images.

However, this tool has some limitations. Probably the most restricting is the computation time. A real time application should be implemented in some other time- efficient language such as C/C++ or similar.

1.5 Scope of Thesis

This work proposes a novel algorithm which is based on simple general parameters which are being used in the field of image processing for many different purposes. Its strength lies in the correct combination of these parameters which outshines many previous used techniques.

1.6 Thesis Outline

Chapter 2 tells about the image fusion abstraction level used in the proposed approach. Chapter 3 gives a brief overview of the decision parameters used. Chapter 4 discusses the implementation methodology for the proposed approach. Chapter 5 presents the experimental results on standard test images against the recommended quality metrics and shows comparisons of the results achieved by the proposed scheme with results obtained from previous fusion techniques. Chapter 6 narrates conclusion and future work.

1.7 Summary

This chapter illustrates the need for image fusion and gives an overview of image fusion fundamentals and its different generalization levels. Many fusion techniques i.e. wavelet transform, multiresolution image fusion methods, image pyramids and artificial neural networks based algorithms have been introduced. Finally, scope of the thesis and structure of the thesis is outlined.

Chapter 2

Multifocus Image Fusion

2.1 Multifocus Images

An image is called multifocus when it has some of its parts (regions/objects) in-focus and other parts out of focus.

Multifocus images can be generated using multiple sensors at various distances or a single sensor over different time slots.

Examples of multi-focus images could be found in optical microscopy. Limited depth of field is a common problem with conventional light microscopy. Since the specimen is sectioned by moving the object along the optical axis, portions of the objects' surface outside the optical plane will be defocused in the acquired image.

2.2 Multifocus Image Fusion

The process of combining relevant information (i.e. parts of images which are in-focus or clear) from a set of input images (of the same scene) into a single image.

It can be clearly viewed from Figure 2.1. Here we've two multifocus pictures. First one is left focused i.e. its left part is in-focus (clearly visible) and right half is out of focus (blurred) while second one is right focused i.e. left half is blurred and right half is clear. When we give these two multifocus images to a multifocus image fusion algorithm as input images, it'll give us one clear output image which has both left and right half clear. The technique is to take both clearer parts from input images and leave the blurred ones.

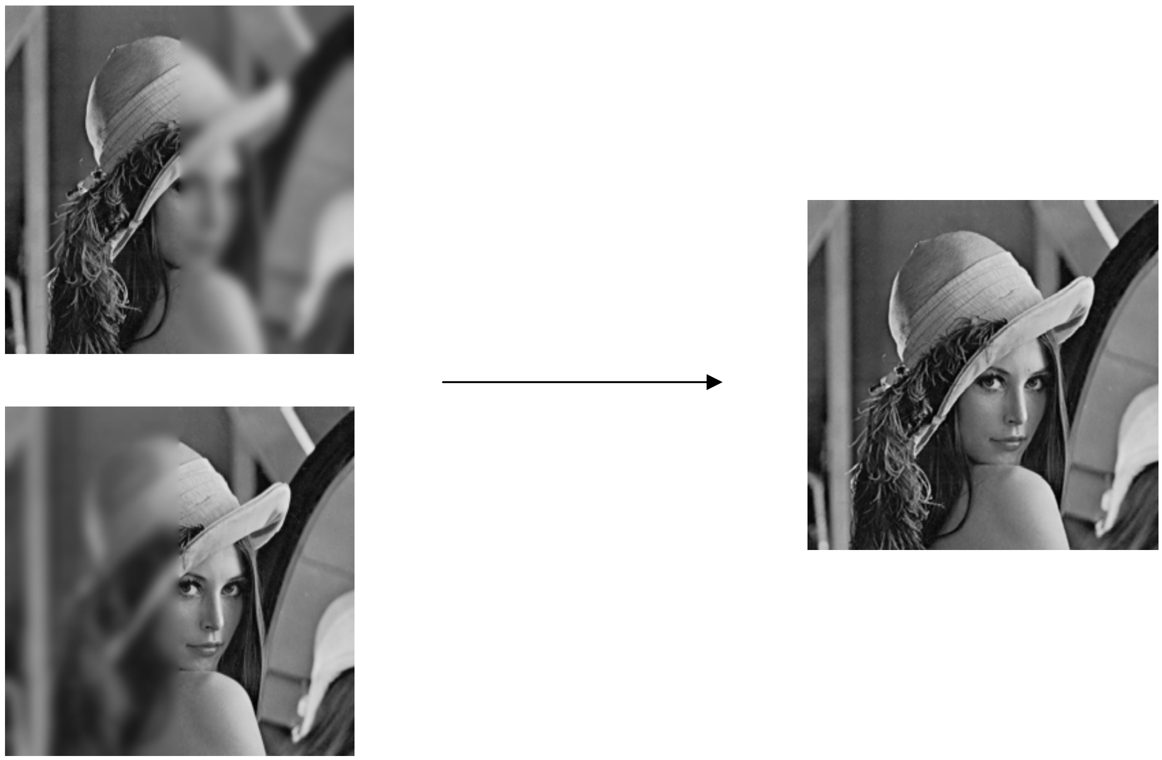


Figure 2.1: Multifocus Image fusion of Lena Image

So the objective required from multifocus image fusion algorithm is to generate a fused image which describes the scene better than any single input image with respect to some relevant properties, by:

- ✓ Extracting all the useful information from the source images, and
- ✓ Not introducing artifacts or inconsistencies which will distract human observers or machine processing.

The advantages of multi-focus images can be fully exploited by merging the sharply focused regions into one image that will be in-focus everywhere.

2.3 Multifocus Image Fusion Methods

Mainly there are three classes of multifocus image fusion methods.

2.3.1 Region Selection Methods

In region selection methods, the input images are initially divided in regions typically blocks or into segments using a segmentation technique. From sets of such regions, one region per set is chosen based on a sharpness criterion and blended or fused to form the final fused image. The value of the sharpness criterion increases and decreases as objects come into focus and go out of focus, or if the contrast changes in the scene. In region based methods, regions are typically selected in the image based on edges. In segmentation based methods, primarily high frequency information in the area of depth of field (DOF) in each input image is used. When the DOF is narrow, edge correspondence between the input images is not the same, due to the optics of the system. Therefore, segmentation based on physical object boundaries becomes ambiguous. In methods using tiling, the most widely reported issues are from blocking effects.

2.3.2 Learning Based Methods

Llearning based methods use training engines which learn to classify between sharp and blurred regions and are normally computationally expensive. Training is normally done with prescribed focused and unfocused training data sets. In the advent of a region that is blurred in all the input images, i.e., unseen data, misclassification takes place and learning based methods employ averaging or force on arbitrary region as the fused image. Ringing effects have been widely reported.

2.3.3 Multi-scale Decomposition (MSD) Methods

In MSD based methods, many of the fusion regulations used, rely on pixel manipulation or replacement at a detail level and, these results in changes in the intensity values of the fused image. These effects are not very prominent to the end user when using simple data sets but stands out as an issue in accurate scene inspection. Furthermore, methods employing wavelets for a particular application may not be extendible to another application as it is difficult to realize a wavelet kernel that can handle multi-scale data sets.

2.4 Advantages of Multifocus Image Fusion

Some potential advantages of multifocus image fusion are:

i. We can make a trust worthy system by employing multiple sensors. If one or two sensors go out of work, still we have other sensors to continue functionality and all the work does not need to be stopped. So sometimes, redundancy can be helpful in the robustness of a system.
ii. If we install multiple sensors and at the same time we have some image fusion mechanism, we do need to store all the information from all sensors. We can just save compact information from all the sensors.
iii. Many different sensors can be used for different atmospheres and time states for example some can work effectively in day light while some other in night. So we can have extended range of operations under different operating conditions. We don't need to worry about composite output because we've image fusion algorithm to take care of.
iv. Different sensors have different spatial and temporal resolutions. So we can use several sensors to obtain joint information. Reduced uncertainty: joint information from multiple sensors can reduce the uncertainty associated with the sensing or decision process.
v. When we use multiple sensors, we can have multiple readings. These multiple readings can reduce the fear of noise or other artifacts in the results. So we can avoid uncertainty in sensitive real time applications.

2.5 Related Research Fields

Here are some related fields where we can use multifocus image fusion.

2.5.1 Intelligent Robots

i. Require motion control, based on feedback from the environment from visual, tactile, force/torque, and other types of sensors
ii. Stereo camera fusion
iii. Intelligent viewing control
iv. Automatic target recognition and tracking

2.5.2 Medical Imaging

i. Fusing X-ray computed tomography (CT) and magnetic resonance (MR) images
ii. Computer assisted surgery
iii. Spatial registration of 3-D surface

2.5.3 Manufacturing

i. Electronic circuit and component inspection
ii. Product surface measurement and inspection
iii. Non-destructive material inspection
iv. Manufacture process monitoring
v. Complex machine/device diagnostics
vi. Intelligent robots on assembly lines

2.5.4 Military and Law Enforcement

i. Detection, tracking, identification of ocean (air, ground) target/event
ii. Concealed weapon detection
iii. Battle-field monitoring
iv. Night pilot guidance

2.5.5 Remote Sensing

i. Using various parts of the electro-magnetic spectrum
ii. Sensors: from black-and-white aerial photography to multi-spectral active microwave space-borne imaging radar
iii. Fusion techniques are classified into photographic method and numerical method

2.6 Summary

This chapter gives a brief overview of what multifocus images are and how multifocus images fusion operate. It tells about three major classes of multifocus image fusion techniques. Some salient advantages of image fusion are outlined and also mentioned some prominent practical fields where image fusion techniques can be employed.

Chapter 3

Feature Level Image Fusion

3.1 Overview of Feature Level Image Fusion

Feature level Image Fusion is also called Intermediate level Image Fusion. It is in between basic Pixel level Image fusion and Decision based Image Fusion. In this type, an image is divided into partitions on multiple levels. These partitions can also be called regions, so its other name is Region level Image Fusion. Each of these regions represents a sub portion of the whole image. Since Image Fusion is aimed to fuse two or multiple images in which some parts are blurred while other are clear, that's why it is beneficial to partition these input images into regions. This practice can differentiate good and bad parts of images into separate categories because after partitioning every blurred or clear part must lie in some region of partitioned image. There may be possibility that some region can have both blurred and clear parts. To overcome this, that region is again subdivided into much smaller part. After a sequence of this partition mechanism, one can assure that each good and bad part resides in a different portion and image has been properly segmented. After such partitioning, several rules can be applied to differentiate between blurred and clear parts of image. These rules are called decision parameters or features. These features clearly tell about health of different regions or segments of image. In most applications of image fusion, people pay more attention to fused objects rather than individual pixels. Since objects could be represented by regions, thus region-based fusion approaches could be more meaningful than pixel-based fusion methods.

There is a compromise on the size of these partitions or regions. If we take a partition too big, there may be a possibility that this bigger portion can have many blurred and clear parts contained in it. On the other hand, if we take a region or partition too small, computation complexity increases. We've to calculate decision parameters for each of these small portions which is unnecessary. So we've to be very careful about choosing the order of partition. It can be compromised taking in consideration both issues in mind.

A number of region-based fusion schemes have been proposed. These methods initially transform the pre-registered images to multi-scale representations, and regions are extracted from each source image. The source images are then fused based on a simple region property such as average activity. These methods do not take full advantage of the wealth of information that can be calculated for each region.

3.2 Decision features extractions methods

In region level image fusion (also named as Feature level image fusion) combines various features. These features may come from several raw data sources (several sensors, different moments, etc.) or from the same raw data. In the latter case, the objective is to find relevant features among available features that might come from several feature extraction methods. The objective is to obtain a limited number of relevant features. Typically, in image processing, feature maps are computed as pre-processing for segmentation or detection. Features such as edges, corners, lines, texture parameters (Haralick, Wavelet coefficients, etc.) are computed and combined in a fused feature map that may then be used for segmentation or detection. General operations for extracting features from images partitioned into regions are:

- ✓ Geometric operations
- ✓ Neighborhood and block operations
- ✓ Transforms
- ✓ Image analysis and enhancement
- ✓ Binary image operations
- ✓ Region of interest operations

3.3 Transforms

Transforms are no doubt the most popular and effective operations used in feature extraction techniques. Different transform functions are used to extract useful features from images and these features are then used in some other algorithms like algorithms for image fusion. Among different transforms, wavelet transform is most widely used and effective transform and it also has been used in many feature extraction methods as building block.

Although the Fourier transform has been the foundation of transform-based image processing since the late 1950s, but it is well suited only to the study of stationary data where all frequencies have an infinite coherence time. The Fourier transform brings only global information which is not sufficient to detect compact patterns in images. Gabor introduced a local Fourier transform, taking into account a sliding window, leading to a time frequency-analysis. This method is only applicable to situations where the coherence time is independent of the frequency. Morlet introduced the wavelet transform in order to have a coherence time proportional to the period.

3.3.1 Overview of Wavelet Transform

Images are generally connected regions of similar texture and gray level that combine to form objects. If the objects are small in size or low in contrast, they are examined at high resolutions; if they are large in size or high in contrast, a coarse view is all that is required. If both small and large objects (or low and high contrast objects) are present simultaneously, it can be advantageous to study them at several resolutions. Wavelet transform is concerned with the representation and analysis of images at more than one resolution.

Wavelet transform represents a windowing technique with variable-sized regions. Wavelet transform allows the use of long time intervals where we want more precise low-frequency information, and shorter regions where we want high-frequency information. Wavelet analysis is capable of revealing aspects of data (e.g., trends, breakdown points, discontinuities in higher derivatives, and self-similarity) that other image/signal processing techniques miss.

3.3.2 What is a Wavelet?

A wavelet is a waveform of effectively limited duration that has an average value of zero. Sinusoids, which are the basis of Fourier analysis, do not have limited duration; they extend from minus to plus infinity. Also, sinusoids are smooth and predictable, whereas wavelets tend to be irregular and asymmetric. This comparison of sine wave and a wavelet is shown in Figure 3.1.

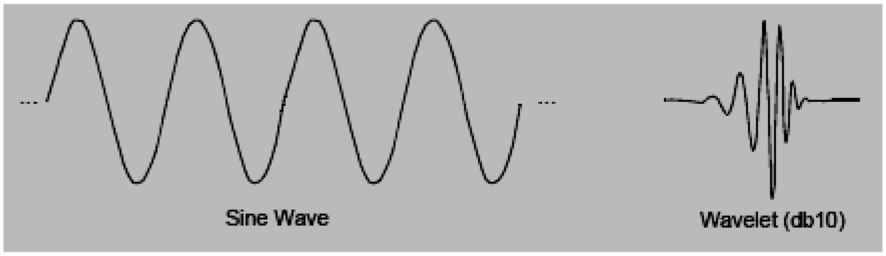


Figure 3.1: Comparison of a Sine Wave and a Wavelet [34]

Fourier analysis consists of breaking up a signal into sine waves of various frequencies. Similarly, wavelet analysis is the breaking up of a signal into shifted and scaled versions of the original (or mother) wavelet.

Just looking at pictures of wavelets and sine waves in Figure 3.1, it is evident that signals with sharp changes might be better analyzed with an irregular wavelet than with a smooth sinusoid. It also makes sense that local features can be described better with wavelets that have local extent.

3.3.2.1 Scaling

Scaling a wavelet simply means stretching (or compressing) it. To describe the effect of scaling, a new term, the scale factor, has been introduced, often denoted by the letter a. The smaller the scale factor, the more "compressed" the wavelet. The Figure 3.2 shows three versions of a wavelet with three different values of the scale factor. As the scale factor decreases, the wavelet is more compressed.

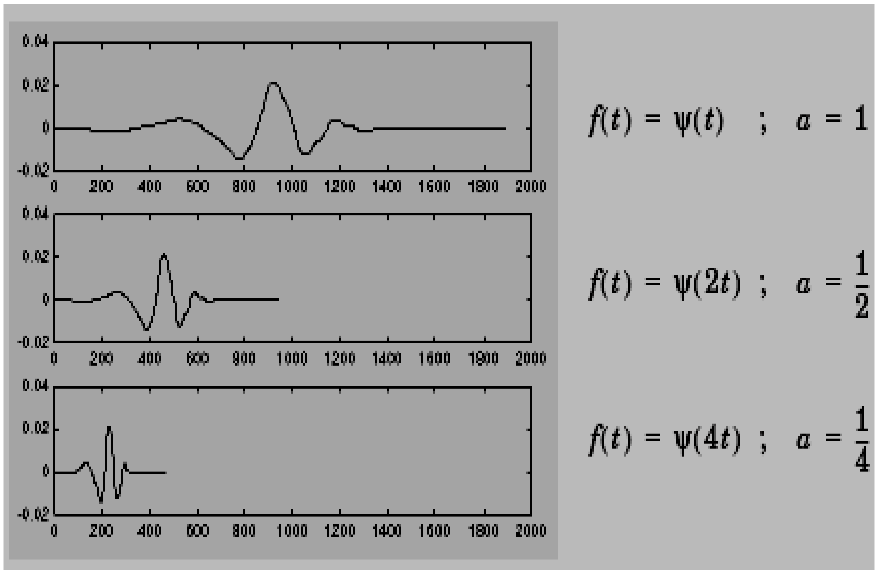


Figure 3.2: Effect of scale factor on a wavelet [34]

3.3.2.2 Shifting

Shifting a wavelet simply means delaying (or hastening) its onset. Mathematically, delaying a function ψ (t) by k is represented by ψ (t-k). It is graphically shown in Figure 3.3 below.

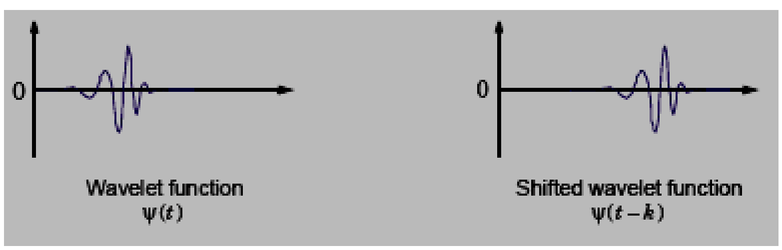


Figure 3.3: Effect of shifting a wavelet [34]

3.3.2.3 Scale and Frequency

As the higher scales correspond to the most "stretched" wavelets, the more stretched the wavelet, the longer is the portion of the signal with which it is being compared, and thus the coarser are the signal features being measured by the wavelet coefficients as depicted in the Figure 3.4.

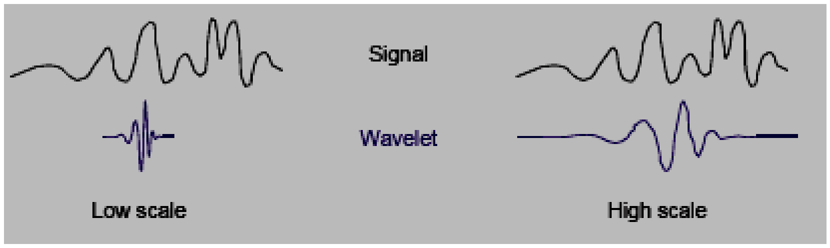


Figure 3.4: Relationship between scale and frequency of a wavelet [34]

Thus, there is a correspondence between wavelet scales and frequency as revealed by wavelet analysis:

Low scale $\Rightarrow$ Compressed wavelet $\Rightarrow$ Rapidly changing details $\Rightarrow$ High frequency.

High scale $\Rightarrow$ Stretched wavelet $\Rightarrow$ Slowly changing features $\Rightarrow$ Low frequency.

3.3.3 One-dimensional Discrete Wavelet Transform

The discrete wavelet transform (DWT) of a function $f(x)$ is given by the following pair of equations [21].

$$W_{\varphi}(j_0,k) = \frac{1}{\sqrt{M}} \sum_{x} f(x)\varphi_{j_0,k}(x) \tag{1}$$

$$W_{\psi}(j,k) = \frac{1}{\sqrt{M}} \sum_{x} f(x)\psi_{j,k}(x) \tag{2}$$

for $j \geq j_0$ and

$$f(x) = \frac{1}{\sqrt{M}} \sum_{k} W_{\varphi}(j_0,k)\varphi_{j_0,k}(x) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_{k} W_{\psi}(j,k)\psi_{j,k}(x) \tag{3}$$

Here $f(x)$, $\varphi_{j_0,k}(x)$ and $\psi_{j,k}(x)$ are functions of the discrete variable x = 0, 1, 2, …, M-1. Normally $j_0 = 0$ and M is a power of 2. The coefficients defined in Eqs. (2) and (3) are called approximation and detail coefficients respectively.

An efficient algorithm to implement DWT using filters was developed in 1988 by Mallat. This very practical filtering algorithm yields a fast wavelet transform (FWT) – a box into which a signal passes, and out of which wavelet coefficients quickly emerge.

For many signals, the low-frequency content is the most important part. It is what gives the signal its identity. The high-frequency content, on the other hand, imparts flavor or gradation. In wavelet analysis, the approximations are the high-scale, low-frequency components of the signal and the details are the low-scale, high-frequency components. The filtering process, at its most basic level, looks like as displayed in Figure 3.5:

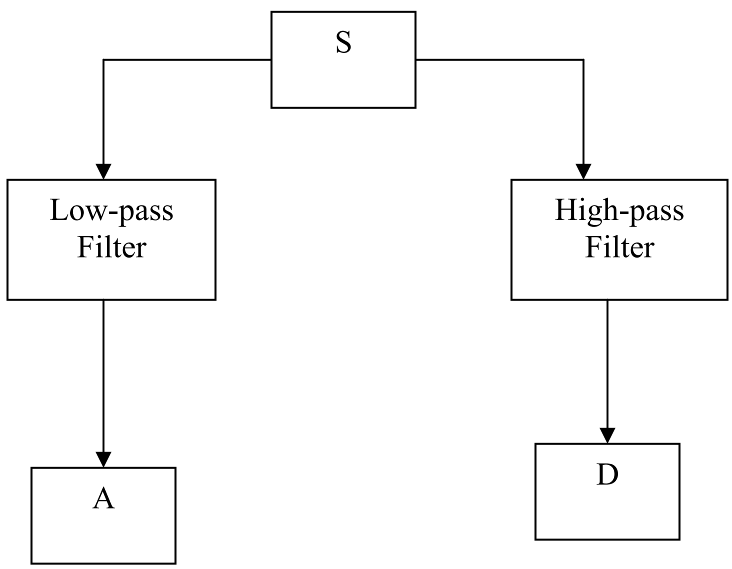


Figure 3.5: An FWT analysis bank

The original signal, S, passes through two complementary filters and emerges as two signals A(approximation) and D(detailed). If we actually perform this operation on a real digital signal, we wind up with twice as much data as we started with. Suppose, for instance, that the original signal S consists of 1000 samples of data. Then the resulting signals will each have 1000 samples, for a total of 2000. But there exists a more delicate way to perform the decomposition using wavelets. By looking carefully at the computation, we may keep only one point out of two in each of the two 2000-length samples to get the complete information.

The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components. The Figure 3.6 shows decomposition of the signal S at the third level.

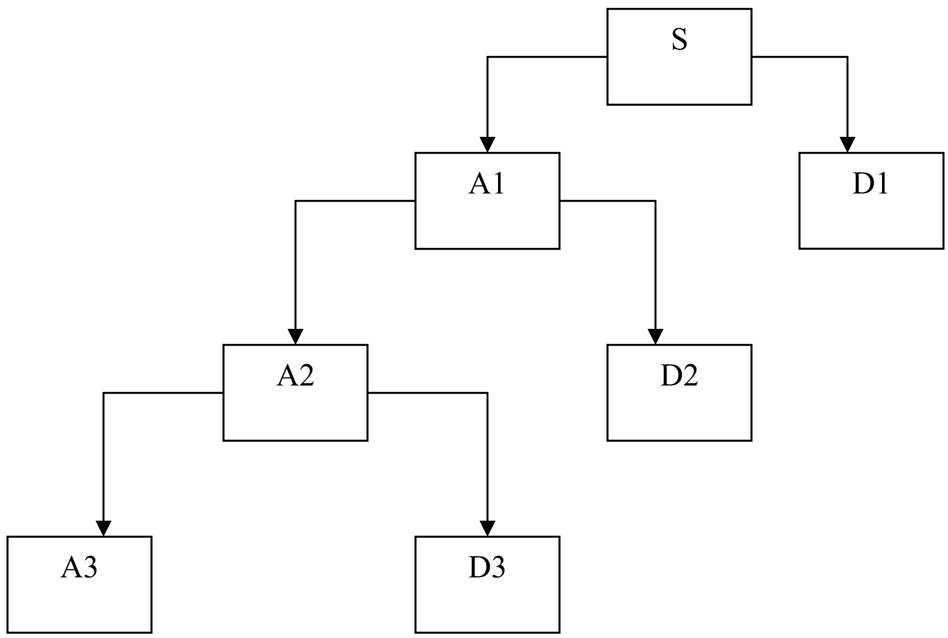


Figure 3.6: Multilevel wavelet-decomposition of a signal

3.3.4 Two-dimensional Discrete Wavelet Transform

In two dimensions, a two-dimensional scaling function, $\varphi(x, y)$ and three two dimensional wavelets $\psi^H(x, y)$, $\psi^V(x, y)$ and $\psi^D(x, y)$ are required [33]. Each is the product of a one-dimensional scaling function φ and corresponding wavelet ψ. Excluding products that produce one-dimensional results, like $\varphi(x)\varphi(x)$, the four remaining products produce the separable scaling function

$$\varphi(x, y) = \varphi(x)\varphi(y) \tag{4}$$

and separable, "directionally sensitive" wavelets

$$\psi^H(x, y) = \psi(x)\varphi(y) \tag{5}$$

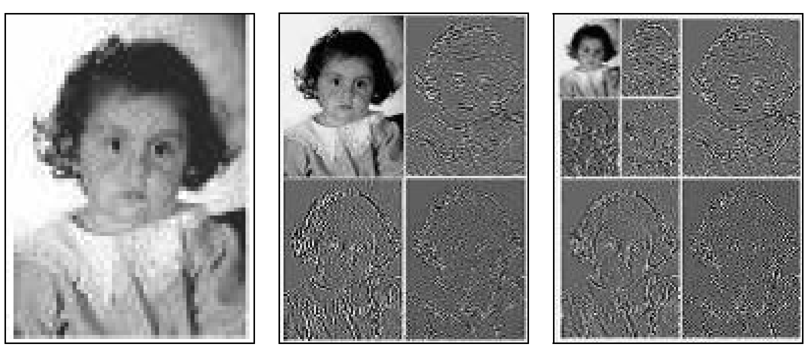
$$\psi^{V}(x,y)=\varphi(x)\psi(y) \tag{6}$$

$$\psi^{D}(x,y)=\psi(x)\psi(y) \tag{7}$$

These wavelets measure functional variations (i.e., intensity or gray-level variations for images) along different directions: ψ^{H} measures horizontal variations, ψ^{V} responds to vertical variations, and ψ^{D} corresponds to variations along diagonals.

Like the one-dimensional discrete wavelet transform, the two-dimensional discrete wavelet transform can be implemented using filters and downsamplers. With separable two-dimensional scaling and wavelet functions, we simply take the one-dimensional FWT of the rows of the two dimensional function $f(x,y)$, followed by the one-dimensional FWT of the resulting columns. Like its one-dimensional counterpart, the two-dimensional FWT filters the scale j + 1 approximation coefficients to construct the scale j approximation and detail coefficients. In the two-dimensional case, however, we get three detail coefficients – the horizontal, vertical and diagonal details.

The subimages, which are shown in Figure 3.7(b), are the inner products of the image in Figure 3.7(a) and the two-dimensional scaling and wavelet functions in Eqs. (2.4-1) through (2.4-4), followed by downsampling by two in each dimension. Two iterations of the filtering process produce the two-scale decomposition as shown in Figure 3.7(c).



(a) (b) (c)

Figure 3.7: The resulting decomposition of a two-dimensional FWT [35]

3.4 Advantages of Feature level Image Fusion

Feature level image fusion or region-based fusion scheme has following advantages.

- ✓ It has reduced sensitivity to noise.
- ✓ It has less blurring effects.
- ✓ It also reduces problems of mis-registration.

3.5 Summary

This chapter deals with abstraction level of image fusion i.e. Feature level image fusion which is also being adopted in this work. In this technique, images are partitioned into regions and decision features are extracted from these regions. Some conventional methods for feature extractions are pointed out. At the end, some benefits of region based image fusion are stated.

Chapter 4

Decision Features

4.1 Overview

Decision features are defined as those distinct distinguishing parameters which help in taking decisions about some issue and maintain some criteria on the basis of which one can distinguish among different sets of things. As stated in previous chapter, in region based image fusion, input images are partitioned into regions. After this, some decision features are extracted from these regions of each input image. Then some criteria about these features are maintained on the basis of which it is decided to pick some regions from one image and other from some other images.

These features are based on the requirements of the problem we've to solve. These features generally, describe overall condition of the concerned region of the image from some aspect. In the fields of image fusion, these features tell about physical health of the region i.e. whether that region is clear or blurred. These parameters are based on physical properties of pixels of the images which tell about condition of the concerned region. A region will be clearer if it is coarser. Similarly some features looks for gray scale variations. If the variation is high, the more will be spatial frequency of that portion of the image, and clearer it will be. If variations in gray scale are less, than it will be less clear. On extreme, if there is no variation at all, it means the image portion is completely blurred.

Features like edges, corners, lines, texture can also be used for distinguishing different regions of image. After decision about each region of an image, a binary decision map is generated for each input image. These maps provide a fruitful observation of health of all regions of the image. On the basis of these map, we can easily determine which part of image is clear and which is blurred. Through this practice, we can pick clear instances of regions from different images and reconstruct a better image than all input images.

4.2 Decision Parameters used in proposed approach

In the proposed approach described below, three parameters are used which decide about the clarity of images partitioned into blocks. These parameters tell which block from which input multi-focus images are more suitable to pick and choose. These parameters are used for blocks, not for whole image. These are Image Clarity [6], Spatial Frequency [7], and Block Visibility [8].

4.2.1 Image Clarity

Image clarity (IC) is a measure which has strong ability to distinguish the clear parts of an image from the blurred parts. In multi-focus image fusion, the clearer the original image block is thought to be, the coarser the original image block is and thus the higher value of IC it has.

$$IC = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{\left(\frac{\partial f(i,j)}{\partial x}\right)^2 + \left(\frac{\partial f(i,j)}{\partial y}\right)^2} \qquad (8)$$

Here, f is an input block to calculate IC and M and N represent the size of the image

4.2.2 Spatial Frequency

Spatial frequency represents the manner in which gray-scale values change relative to their neighbors within an image. If there is a slowly varying change in gray scale in an image from one side of the image to the other, the image is said to have a low spatial frequency. If pixel values vary radically for adjacent pixels in an image, the image is said to have a high spatial frequency. It has been observed on experimental basis that parts of an image which are clearer always have higher SF than the blurred blocks. If input image is f, then SF is:

$$SF = \sqrt{RF^2 + CF^2} \qquad (9)$$

$$RF = \sqrt{\frac{1}{(M-1)(N-1)}\sum_{i=1}^{M-1}\sum_{j=1}^{N-1}(f(i,j+1)-f(i,j))^2} \qquad (10)$$

$$CF = \sqrt{\frac{1}{(M-1)(N-1)}\sum_{i=1}^{M-1}\sum_{J=1}^{N-1}(f(i+1,j)-f(i,j))^2} \qquad (11)$$

4.2.3 Block Visibility

Image's block visibility is inspired by human visual system. It tells how much clear a block of an image is to a human being. Image's Block Visibility (VI) for k^{th} input image's block B_k is defined as:

$$VI(B_k) = \frac{1}{mn} \sum_{(x,y)\in B_k} \varnothing(m_k) \frac{|f(x,y) - m_k|}{m_k} \qquad (12)$$

Where B_k is an mxn block of input image,

$$m_k = \frac{1}{mn} \sum_{(x,y)\in B_k} f(x,y) \qquad (13)$$

$$\varnothing(m_k) = \left(\frac{1}{m_k}\right)^{\sigma} \qquad (14)$$

σ is a visual constant ranging from 0.6 to 0.7. In the proposed approach σ is chosen as 0.7.

4.3 Summary

This chapter tells what are decision features and for what purpose those used. It throws light on how to extract these features and what properties can be used to extract features which are useful for image fusion techniques. This chapter also includes a brief introduction of features which are being used in the proposed technique.

Chapter 5

Implementation Methodology

This chapter deals with the design and implementation of proposed approach. MATLAB® has been chosen as a development tool because of the availability of required toolboxes and functions. The simulator is developed, compiled and tested in MATLAB®7.0.

5.1 Design of System

The block diagram of proposed scheme is shown in Figure 5.1.

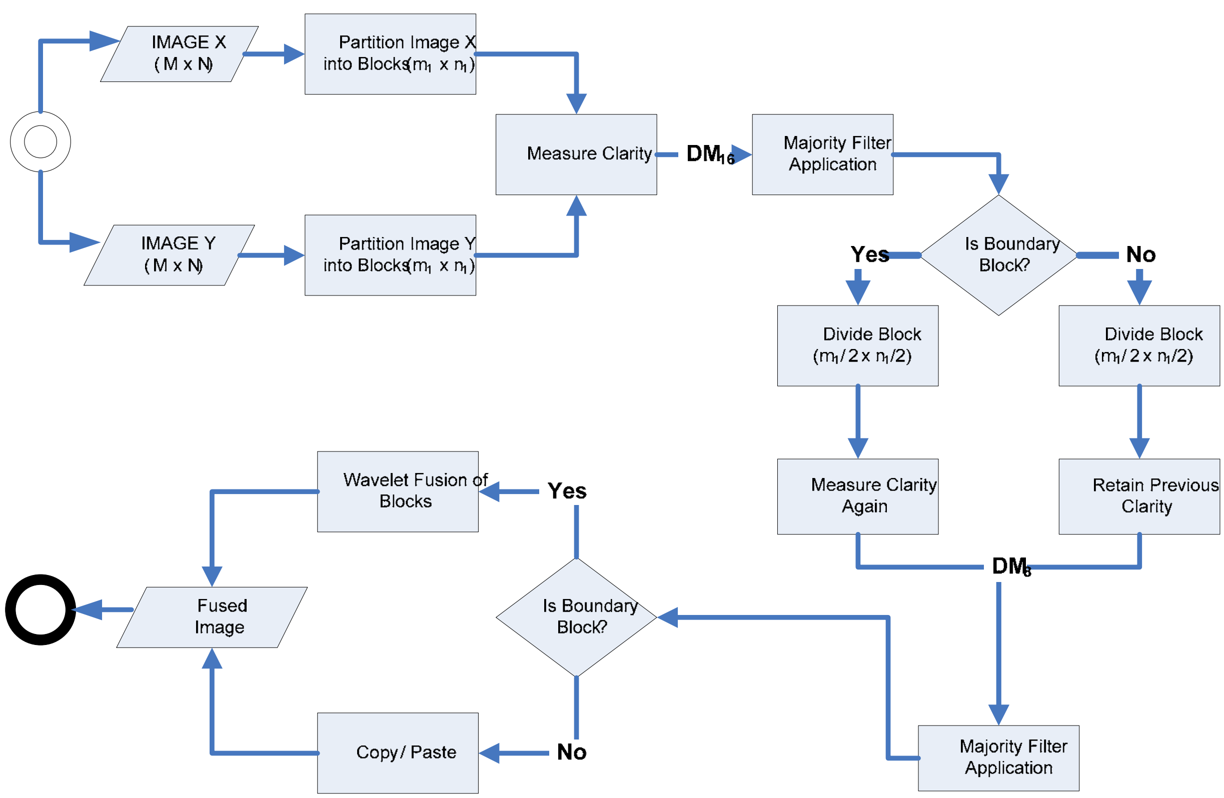


Figure 5.1 Flow Chart of Proposed Scheme

In figure shown above, Images X and Y are two input images of same order MxN and Fused Image is reconstructed image also having same order MxN. All parallelograms represent inputs and output of the system. Rectangles represent operations performed on intermediate inputs. Diamonds represent decisions and arrows show data flow from one construct to another and text on arrow represents intermediate inputs and outputs.

<u>Pseudo-Code of Proposed Approach</u>

The algorithm is follows:

1. Divide each input image i.e. Image A and Image B into blocks of order m_1xn_1. Here m_1xn_1 is 16x16.
2. Make decision map DM_{16} using all three lucidity measuring features i.e. Image Clarity (IC), Block Visibility (BV) and Spatial Frequency (SF), of each block of input images. The criteria would be
   (Any two values of IC_A, BV_A and SF_A) > (Any two values of IC_B, BV_B and SF_B)
3. Apply Majority Filter on DM_{16}.
4. Determine whether any block of DM_{16} is a boundary block.
5. Again divide input images i.e. A and B into block of order m_2xn_2. Here m_2xn_2 is 8x8.
6. Repeat step 2 for each block of partitioned images obtained from step 5, only for those blocks for which their representative value in DM_{16} shows that they are boundary blocks and repeat value of representative element of DM_{16} for all other non boundary blocks. In this way, Decision Map DM_8 is obtained which is four times bigger than DM_{16}.
7. Repeat step 3 for DM_8.
8. Repeat Step 4 for DM_8.
9. Now fuse all boundary blocks by conventional wavelet image fusion technique and copy paste all other blocks from original input images. In this way, final fused image is obtained which is far better than any of the original input image.

5.1.1 Description of proposed scheme

The heart of this scheme is to make a difference between clear and blurred part of input images and on the basis of this decision, make a decision map which clearly tells which block is clearer in which input image. Then using this decision map, output image is reconstructed using multi-focus input images. Here an assumption is made that all the input images are registered and block size is taken as fraction of the order of the input images. For simplicity, two input images are taken so that the results can easily be shown. The same procedure can be done with more than two input images.

Step 1- Image Acquisition

In first step, a grayscale image of order 512x512 is acquired from the World Wide Web to be processed for the fusion process. Image shown in Figure 5.2 represents a standard 'Lena' image of order 512x512.

Figure 5.2: Original Lena test image of size 512x512

Step 2- Application of Average Filter

In second step, two multi-focused images are obtained after applying average filter on original standard input test image. Here we take three options of multi-focusing.:

a) Left and right multi-focused
b) Upper and lower multi-focused
c) Inner and outer multi-focused

In option (a), left portion of first output image is blurred after applying average filter for a number of times and right portion is kept intact. Opposite is done for second image.

In option (b), upper part of first image is blurred while lower portion is kept intact and lower portion of second image is blurred and upper portion is kept intact.

In option (c), center portion of first image is blurred and outer portion of second image is blurred.

The Figure 5.3 shows all these three types of multi-focused images for the original Lena test image shown in Figure 5.2.

The purpose of application of average filter is to get manually generated multi-focused images because correct multi-focused images are not available on World Wide Web. The multi-focused images available on internet don't have their reference image. So for purpose of comparisons with other techniques, it is mandatory to have correct multi-focused images and their reference standard image.

Figure 5.3: Multi-focused images generated using average filter from Lena test image

Step 3- Partitioning both multi-focused images into blocks of order 16x16

In Step 3, both multi-focused images are partitioned into blocks of order 16x16. It is necessary that the order of each partitioned block should be even fraction of the order of original image. In this way, all the image can be divided into blocks and no row or column remain outside of a block.

Step 4- Measuring Clarity

Then for each block of order 16x16, all three decision parameters, explained in above chapter, are calculated. The same process is repeated with all blocks of both input images.

After this process, the values of these three parameters are compared for each corresponding block of both images. The block having any two parameters greater than those of the corresponding block of other image is chosen as clear one. In this way a binary decision map DM_{16} is constructed of the order $M/m_1 x M/n_1$ where M is order of original image i.e.512 and $m_1 x n_1$ is order of each block i.e. 16x16. Figure 5.4 shows Decision Map DM_{16} in which each element represents block of order 16x16.

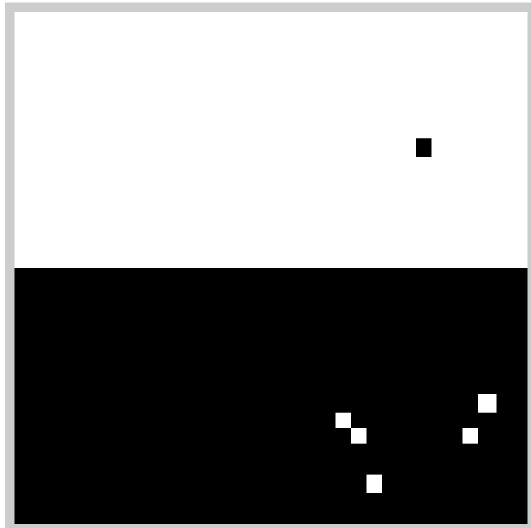


Figure 5.4: Decision Map (DM_{16}) of upper and lower multi-focused Lena Image

It can be observed that there are two distinct parts of DM_{16}. The upper white part represents blocks which should be picked from top focused input image while black part represents blocks which should come from bottom focused input image.

Step 5- Applying Majority Filter on DM_{16}

Figure 5.4 clearly tells that most of the blocks are categorized accurately. But at the same time, some false detected black spots can be seen in white parts and vice versa. It means these falsely detected blocks must lie in other category.

To overcome this false detection, majority filter of order 3x3 is applied on DM_{16} to eliminate small fraction of false detection. So if central element is different from its neighborhood, then it should be rectified. The basic theme of applying majority filter is to include impact of neighborhood blocks of image. Figure 5.5 shows better decision map than Figure 5.4.

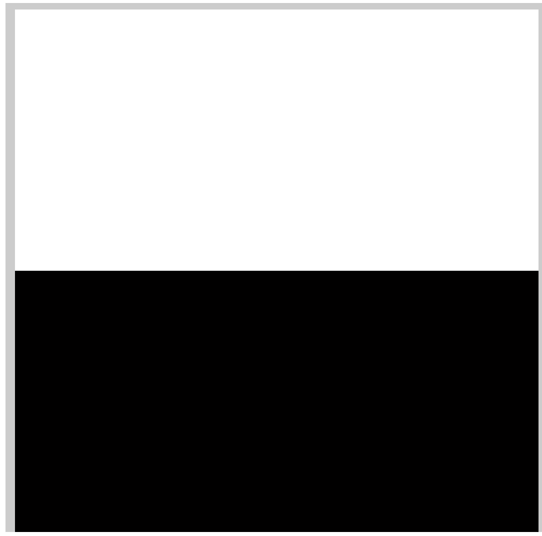


Figure 5.5: Decision Map (DM_{16}) of upper and lower multi-focused Lena Image After applying Majority Filter

Step 6- Is Boundary Block?

After this, the two input images X and Y are again partitioned into blocks of order $m_1/2xn_1/2$ i.e. 8x8. Now using DM_{16}, the elements are found out which represent blocks on the boundary of clear and blurred portions. Three decision parameters are again calculated for these blocks of input images and compared the results. All the other blocks which are not considered as boundary blocks are left. On the basis of this decision, the decision map DM_8 is adjusted. After this practice, the order of DM_8 is doubled the order of DM_{16} because each block represented by each element of DM_{16} is subdivided into four smaller blocks.

Figure 5.6 shows decision map DM_8 in which each element represent block of size 8x8 of input images. It has been built by same decisive technique using decision parameters explained above. Still it can be seen some false detection spots in white part. These falsely detected spots were not identified in DM_{16} because in DM_{16}, each block was of the order of 16x16. It was a bigger block which contained four sub blocks of order 8x8. There may be a possibility that out of these four blocks, one or two blocks belonged to other category. After dividing it into sub blocks, these falsely identified blocks are exposed now.

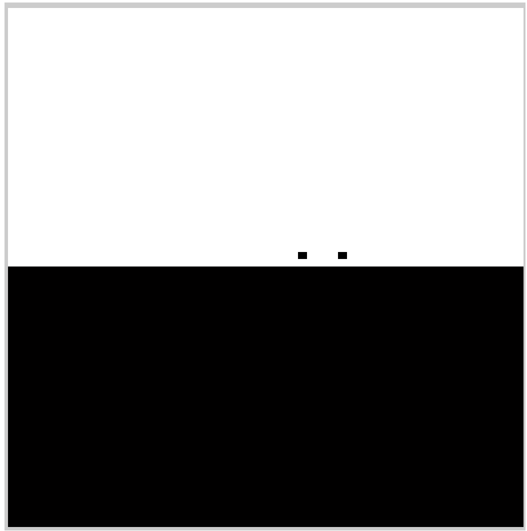


Figure 5.6: Decision Map (DM_8) of upper and lower multi-focused Lena Image

Step 7- Applying Majority Filter on DM8

To eliminate falsely detected blocks in DM8, majority filter of order 3x3 is again run on DM8. Figure 5.7 represents refined DM8 which clearly represents clear blocks from both input images and eliminates blurred blocks.

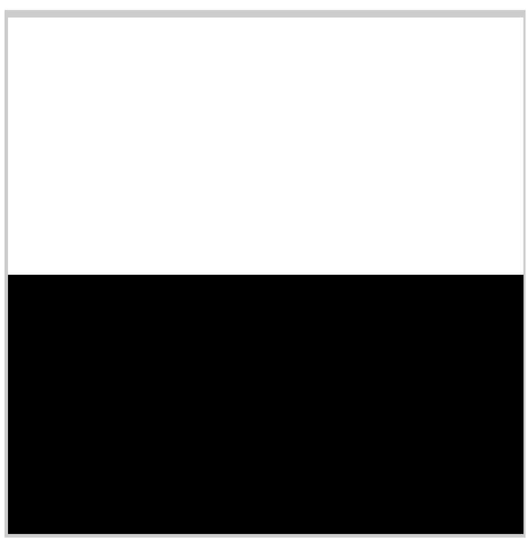


Figure 5.7: Decision Map (DM_8) of upper and lower multi-focused Lena Image After applying Majority Filter

Step 8- Is Boundary Block?

Until now, a correct decision map has been created which clearly divide image into two portions. Upper portion of decision map tells that these blocks must come from image which is blurred from bottom and lower part of decision map tells that blocks representing this part must come from image which is blurred from top. Now bone of contention part is blocks which are boundary of these two parts. These blocks can come from any other part.

So in this step, these boundary locks are identified. This can be done by analyzing neighborhood of these blocks. If number of neighbor blocks from one category is equal to number of blocks from other category, it means these are on boundary. In this way, three parts are clearly identified i.e. one belong to first category, second belong to second category and third belong to boundary blocks.

Step 9 – Reconstruction of fused image

In final step, the blocks of DM_8 which are on boundary on the two regions, are fused using conventional wavelet transform 'db3' at most optimal level '3' while copy paste all other blocks from the input images to which these clear blocks belong.

Figure 5.8 shows fused image which is better than both input multi-focused images.

Figure 5.8: The resulting fused image

5.2 The Simulation

MATLAB® 7.0 has been used for the simulation and implementation of the proposed approach. A screen shot of the simulation is shown in Figure 5.9.

The foremost window of simulation allows the user to browse an image from any place. It supports many formats of image like .bmp, .gif, .jpg etc. There is also a button for reset which will bring to first screen and clears all selections for selecting another image for fusion.

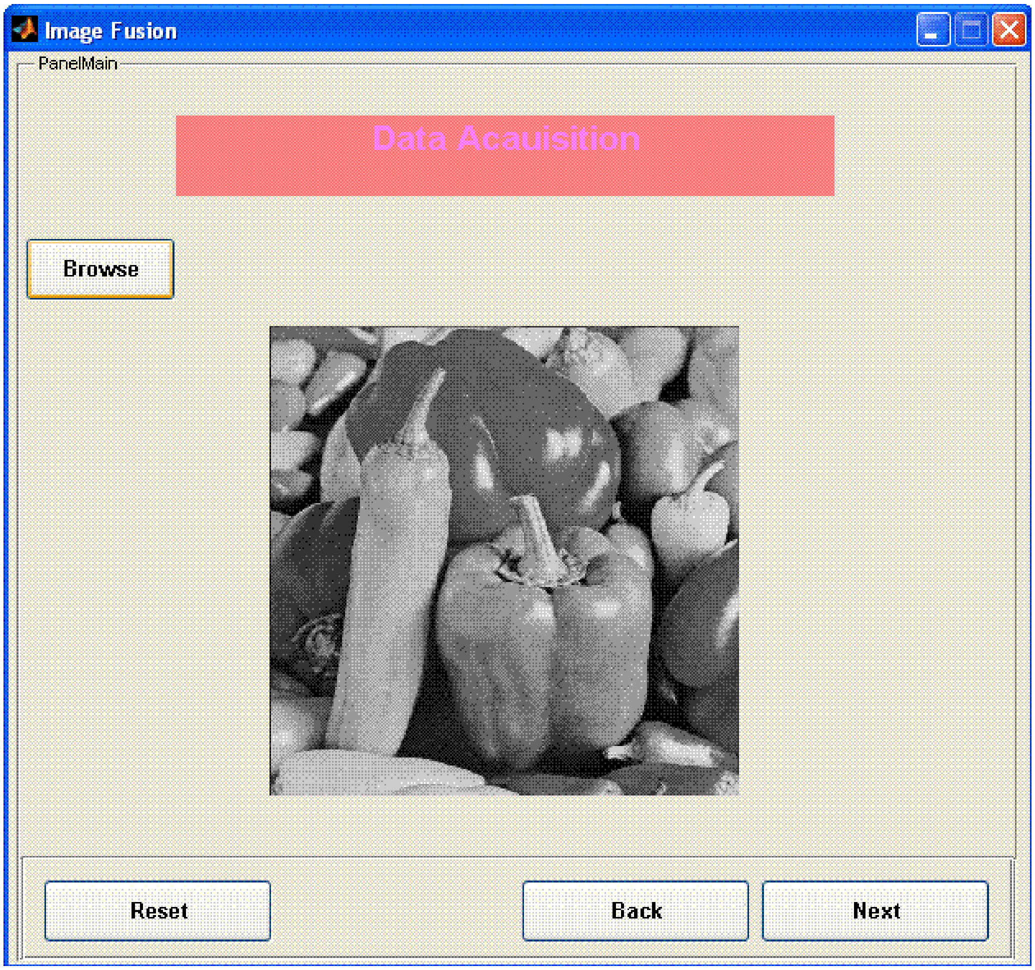


Figure 5.9: Screen shot of Data Acquisition

After browsing an image and pressing Next button, second screen contains multi-focusing options. Here three options are available i.e. left right, top bottom and inner outer. After selecting one option from dropdown list, multi-focused images are displayed below. Average filter has applied for thirty times to get these multi-focused images. There is a button to go back to select another input image as well. Figure 5.10 depicts multi-focusing screen.

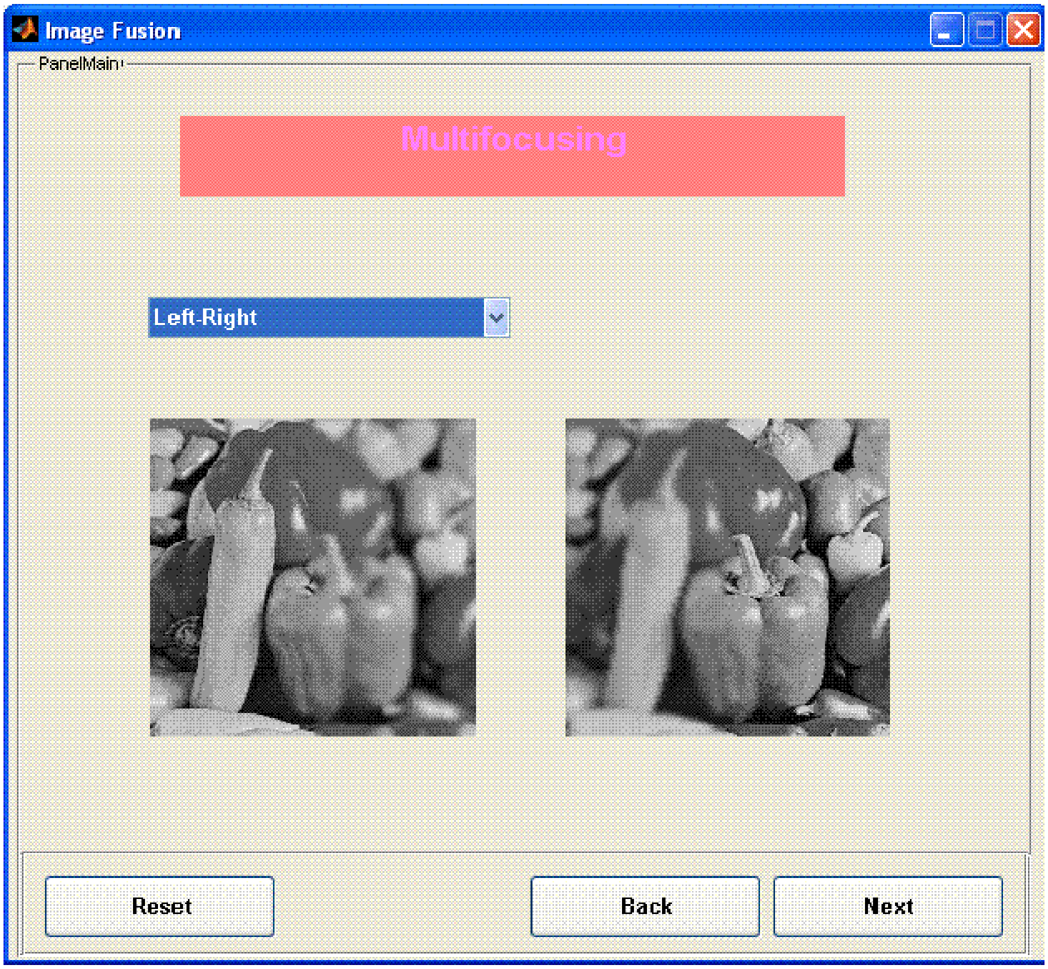


Figure 5.10: Screen shot of Multi-focusing options

After selecting a multi-focusing option and pressing Next button, next screen shows Decision Map16 in which each element represents a block of 16x16 of the original input multi-focused images. Figure 5.11 shows DM16.

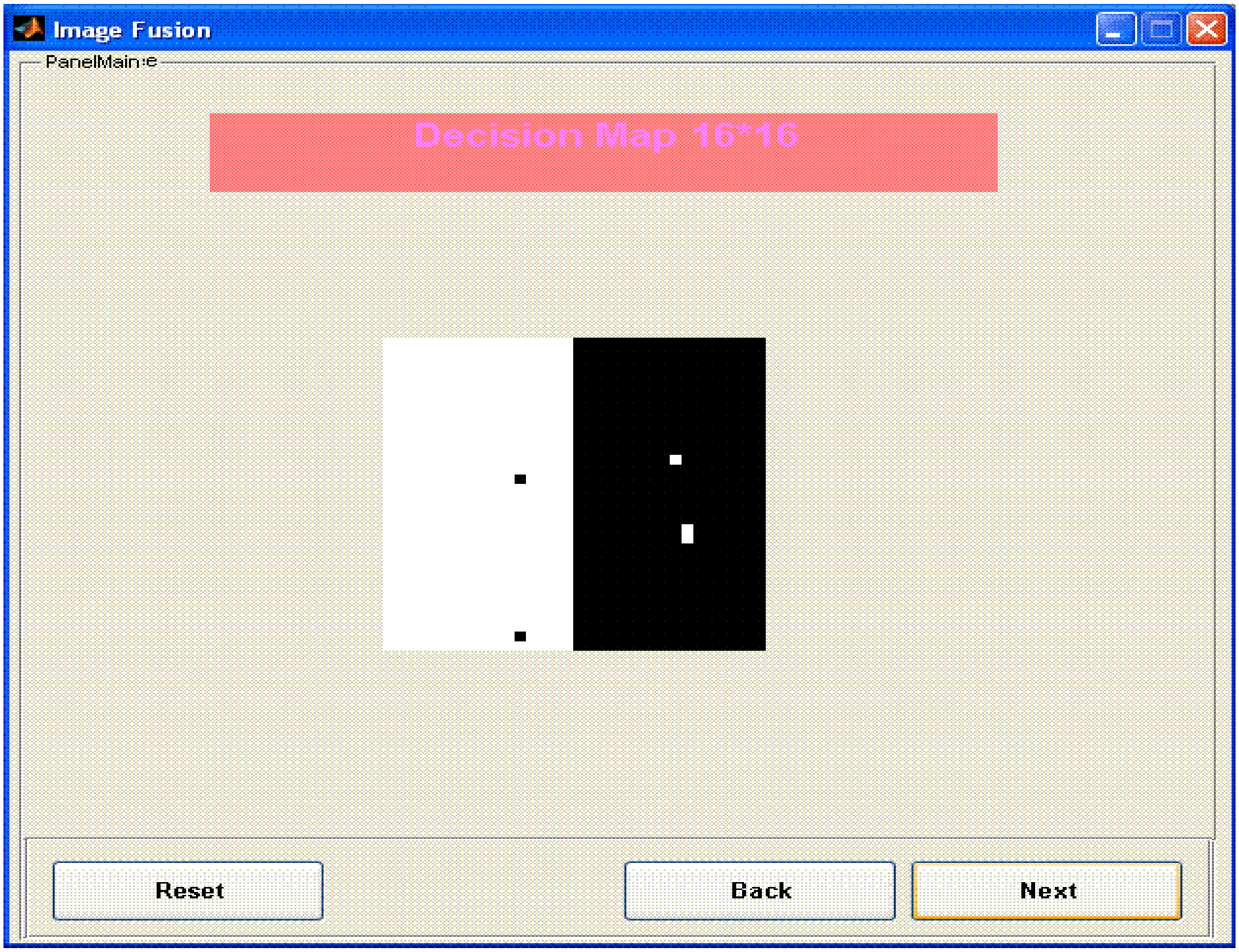


Figure 5.11: Screen shot of Decision Map 16x16

Next screen shows output of applying Majority Filter on Decision Map 16x16.It is also shown in Figure 5.11 below.

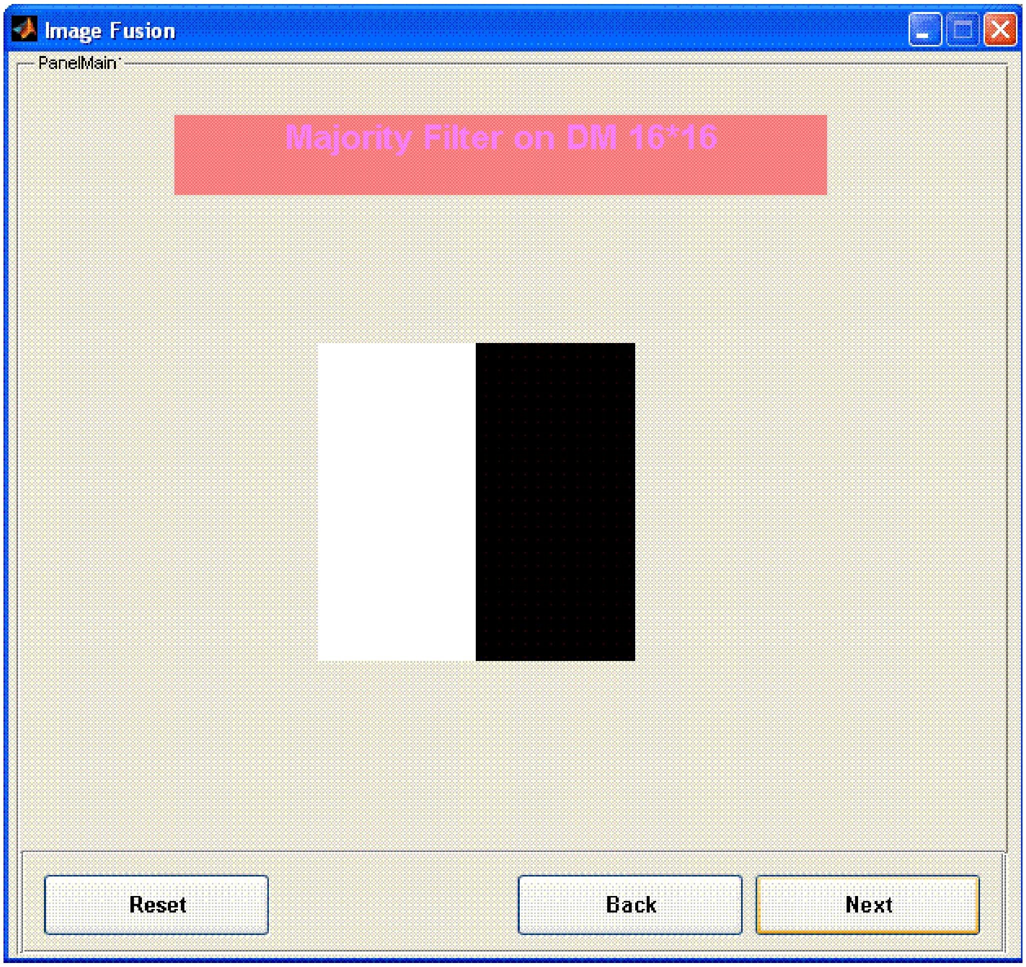


Figure 5.12: Screen shot of Decision Map 16x16 after applying Majority Filter

Figure 5.12 shows screen shot of DM8 in which each element stands for a block of 8x8.

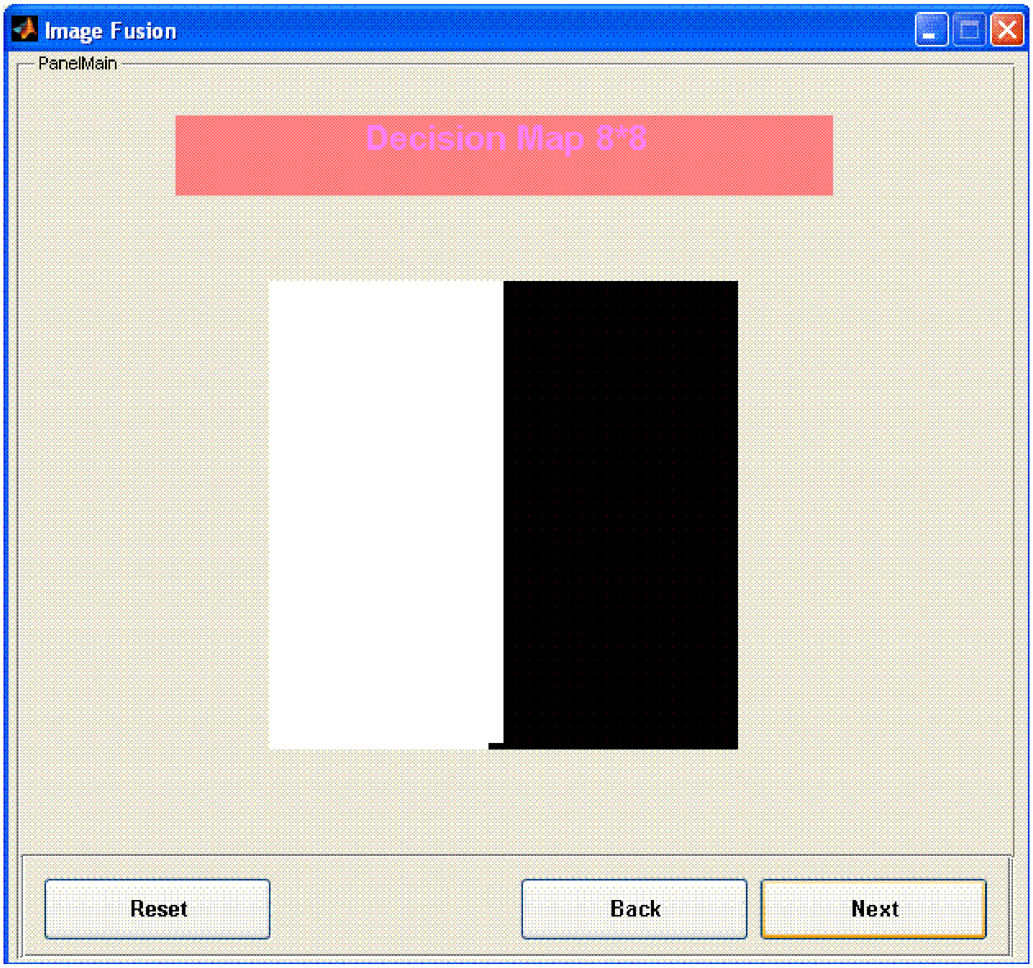


Figure 5.13: Screen shot of Decision Map 8x8

Figure 5.13 shows screen shot of output of applying Majority Filter on Decision Map8.

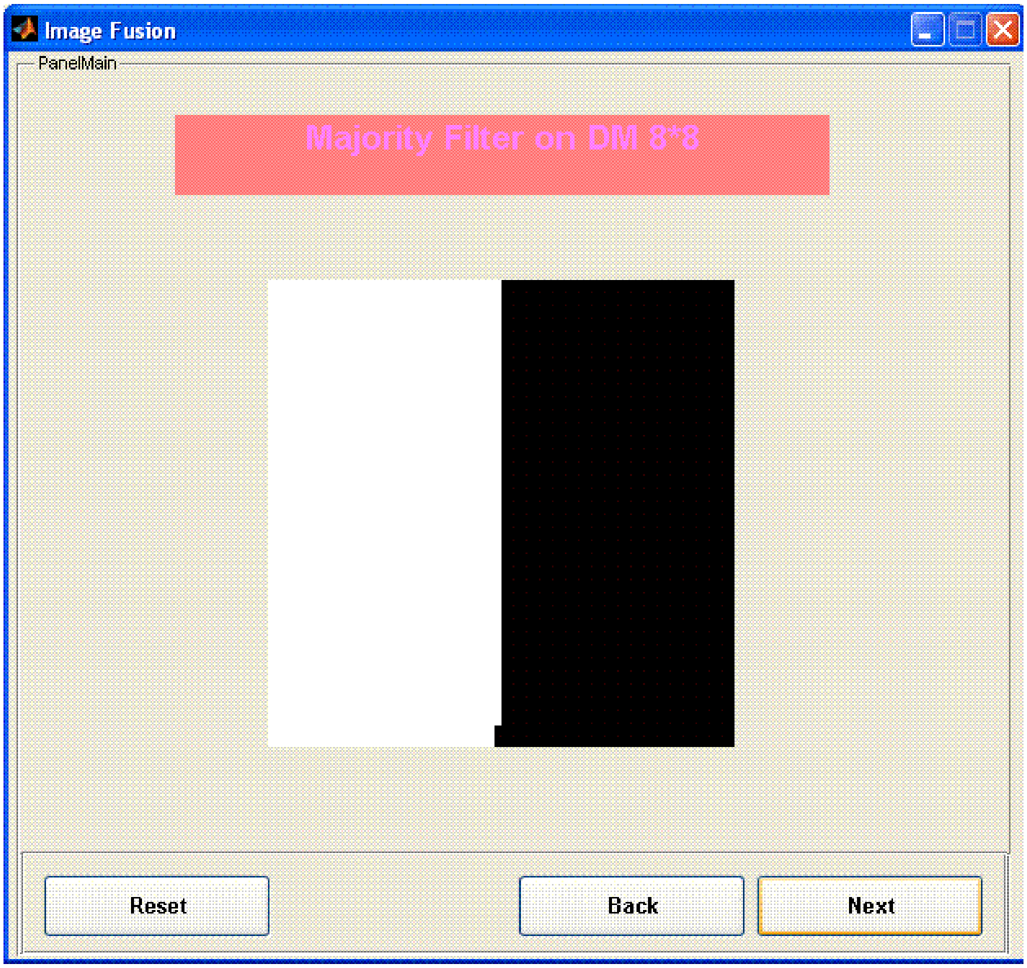


Figure 5.14: Screen shot of Decision 8x8 after applying Majority Filter

Figure 5.14 shows reconstructed image using output of application of Majority Filter on DM8 and input multi-focused images.

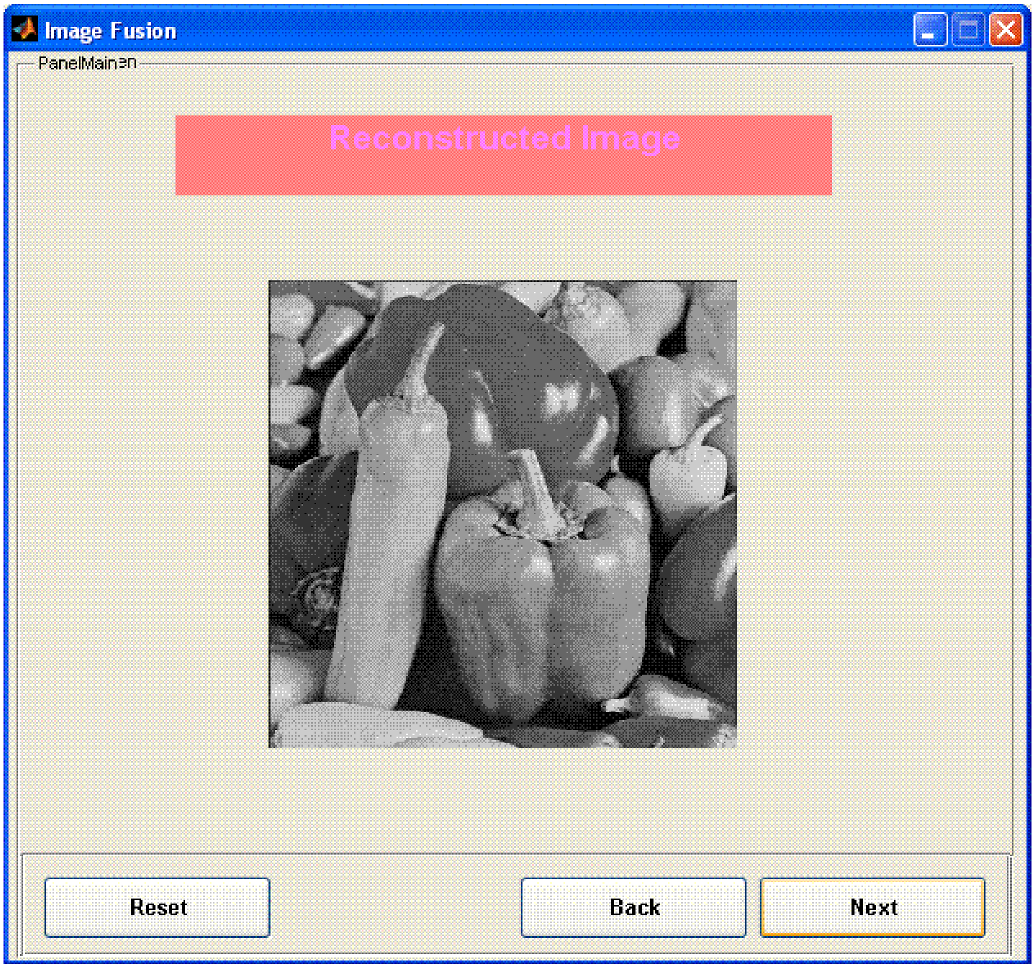


Figure 5.15: Screen shot of Reconstructed Image

Figure 5.15 shows evaluation criteria which show Root Mean Square Error (RMSE) and Peak Signal to Noise Ratio (PSNR) between standard reference image and reconstructed image from our technique.

5.3 Summary

This chapter deals with the entire design of the proposed approach along with its implementation in MATLAB. This chapter also includes graphical display of intermediate results and simulation screen shots of the main implementation.

Chapter 6

Results and Discussion

For comparisons, experiments and discussion, three grayscale standard test images (i.e. Lena, Barbara, and Peppers) of size 512 x 512 have been taken from the World Wide Web. .gif format of these images has been taken because this format yields better results and less computation time than other formats. MATLAB 7.0 has been used for the implementation of the proposed approach and results have been obtained on Pentium-III, 733 MHz processor with a memory of 256 MB. Two different quality metrics i.e. Root Mean Square Error (RMSE) and Peak Signal to Noise Ratio (PSNR) have been used for evaluation and compilation of fusion results. Results have been obtained for different multi-focused regions (i.e. left-right, upper-lower and inner-outer) of the input images.

6.1 Performance Measures

Two types of metrics can be used to measure the performance of fusion system i.e. subjective and objective metrics. The following objective performance metrics have been analyzed to measure the quality of the reconstructed image.

- Peak Signal to Noise Ratio
- Root Mean Square Error

6.1.1 Peak Signal to Noise Ratio (PSNR)

This Qualitative metric tells how much good the reconstructed image is. This objective metric is used to measure the quality of the fused image. More value of PSNR indicate better quality image. For two $M \times N$ monochrome images R and F, where the first one is the original or reference image and the second is the fused image, it is defined as:

$$PSNR = 10.l\,g \frac{255^2 MN}{\sum_{i=1}^{M} \sum_{j=1}^{N} (R(i,j) - F(i,j))^2}. \qquad (16)$$

6.1.2 Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is very useful criteria to measure the difference between two images i.e. reference image and fused image reconstructed by some fusion technique. For a reference image R and the fused image F, both of size $M \times N$, then RMSE is calculated as:

$$RMSE(R,F) = \sqrt{\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[R(i,j) - F(i,j)\right]^{2}}{MN}} \qquad (17)$$

6.2 Experimental Results

In this experiment, three test images are taken (i.e. Lena, Barbara and Peppers) having size 512x512 from World Wide Web. Matlab 7.0 has been used for implementation.

Tables 5.1, 5.2 and 5.3 show the comparisons of fused image and input images of Lena, Peppers and Barbara respectively.

Table 6.1: Comparison of Fused Image wit h Input Images of Lena

Multi-focused Regions	RMSE			PSNR		
	Input Image 1	Input Image 2	Fused Image	Input Image 1	Input Image 2	Fused Image
Left-Right	8.0112	9.0023	**0.4335**	30.0568	29.0438	**55.3912**
Upper-Lower	9.5939	7.2783	**0.7698**	28.4909	30.8903	**50.4031**
Inner-Outer	9.1557	7.7442	**1.2881**	28.8970	30.3512	**45.9318**

Table 6.2: Comparison of Fused Image wit h Input Images of Peppers

Multi-focused Regions	RMSE			PSNR		
	Input Image 1	**Input Image 2**	**Fused Image**	**Input Image 1**	**Input Image 2**	**Fused Image**
Left-Right	11.8229	11.7744	**0.7383**	26.6763	26.7120	**50.7658**
Upper-Lower	10.7791	12.7270	**0.9261**	27.4791	26.0363	**48.7976**
Inner-Outer	14.7310	7.9590	**1.3369**	24.7662	30.1136	**45.6089**

Table 6.3: Comparison of Fused Image wit h Input Images of Barbara

Multi-focused Regions	RMSE			PSNR		
	Input Image 1	**Input Image 2**	**Fused Image**	**Input Image 1**	**Input Image 2**	**Fused Image**
Left-Right	17.2554	12.3477	**0.5147**	23.3923	26.2990	**53.8996**
Upper-Lower	16.0392	13.8631	**0.7595**	24.0271	25.2936	**50.5198**
Inner-Outer	18.7463	9.8957	**1.1643**	22.6725	28.2219	**46.8092**

It becomes obvious after observing results from these tables that fusion process improves the image quality as the RMSE value of fused image is consistently smaller whereas PSNR value is larger than those of the input images.

Table 5.4 through Table 5.6 show comparisons among the proposed approach and some existing methods such as choose maximum wavelet coefficients method (DWT-I), choose maximum wavelet coefficients gradient method (DWT-II), and the methods described in [9-11]. For DWT-based fusion schemes, the wavelet basis "db4" is used. The wavelet decomposition levels of DWT-I and DWT-II are six and five levels that are their own optimal decomposition levels respectively. Consistency verification in a 3x3 window is only used for the DWT-II. The corresponding authors provide results of the schemes in [9-11].

Table 6.4: Image Fusion Methods' Comparisons for Lena Image

Multi-focused Regions	Left and Right		Upper and Lower		Inner and Outer	
Method	**RMSE**	**PSNR**	**RMSE**	**PSNR**	**RMSE**	**PSNR**
DWT-I	1.2983	45.8631	1.4912	44.6603	1.7996	43.0275
DWT-II	1.0285	47.8866	1.4414	44.9552	1.8714	42.6874
[Wen Cao, Bicheng Li, Yong Zhang]	1.3231	45.6991	1.4927	44.6516	1.8198	42.9303
[Zhenhua Li, Zhongliang Jing, Gang Li, Shaoyuan Sun, Henry Leung]	1.1863	46.6466	1.4524	44.8894	1.4282	45.0348
[L. Shutao, W. Yaonan, Z. Changfan]	1.1643	46.8096	1.2434	46.2388	1.3830	45.3141
Proposed	**0.4335**	**55.3912**	**0.7698**	**50.4031**	**1.2881**	**45.9318**

Table 6.5: Image Fusion Methods' Comparisons for Barbara Image

Multi-focused Regions	Left and Right		Upper and Lower		Inner and Outer	
Method	**RMSE**	**PSNR**	**RMSE**	**PSNR**	**RMSE**	**PSNR**
DWT-I	1.8389	42.8398	2.0037	42.0940	2.2532	41.0748
DWT-II	1.3938	45.2469	1.7583	43.2291	2.3986	40.5317
[Wen Cao, Bicheng Li, Yong Zhang]	1.7456	43.2919	1.8919	42.5928	2.1307	41.5603
[Zhenhua Li, Zhongliang Jing, Gang Li, Shaoyuan Sun, Henry Leung]	2.3989	40.5306	2.4298	40.4194	2.5784	39.9039
[L. Shutao, W. Yaonan, Z. Changfan]	1.7834	43.1058	1.8687	42.7001	1.9951	42.1315
Proposed	**0.5147**	**53.8996**	**0.7595**	**50.5198**	**1.1643**	**46.8092**

Table 6.6: Image Fusion Methods' Comparisons for Peppers Image

Multi-focused Regions	Left and Right		Upper and Lower		Inner and Outer	
Method	**RMSE**	**PSNR**	**RMSE**	**PSNR**	**RMSE**	**PSNR**
DWT-I	2.0240	42.0067	2.0930	41.7153	2.3561	40.6871
DWT-II	1.6673	43.6906	1.7773	43.1358	2.1582	41.4490
[Wen Cao, Bicheng Li, Yong Zhang]	2.1073	41.6563	2.1442	41.5053	2.3003	40.8952
[Zhenhua Li, Zhongliang Jing, Gang Li, Shaoyuan Sun, Henry Leung]	5.4568	33.3921	5.4703	33.3706	5.5941	33.1762
[L. Shutao, W. Yaonan, Z. Changfan]	1.8753	42.6694	1.8649	42.7175	1.9694	42.2442
Proposed	**0.7383**	**50.7658**	**0.9261**	**48.7976**	**1.3369**	**45.6089**

Figure 6.1 graphically demonstrate the comparison among RMSE values of proposed scheme and all above referenced techniques for left-right focused images of Lena, Barbara and Peppers. Similarly Figure 6.2 shows comparison among PSNR values of proposed and other techniques.

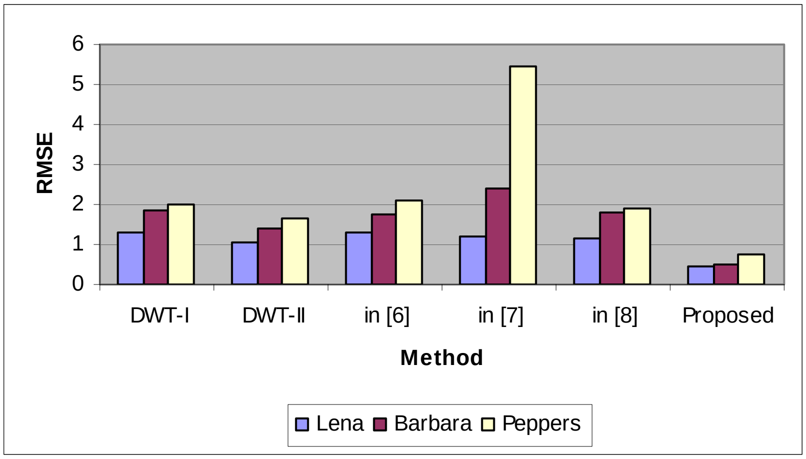


Figure 6.1: Comparison of RMSE

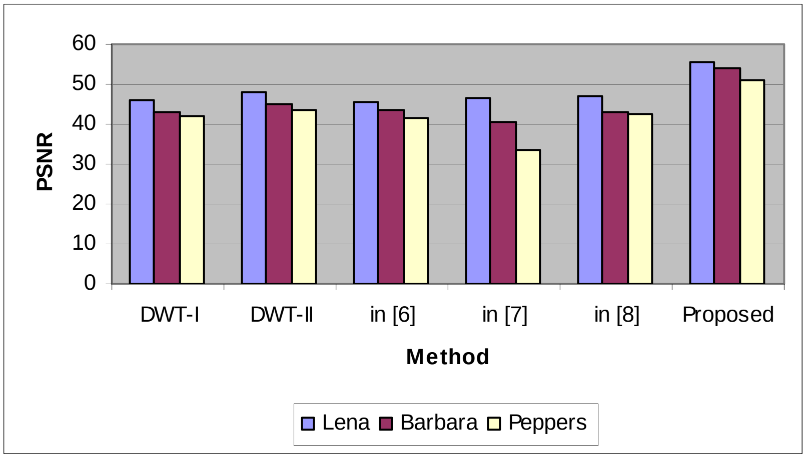


Figure 6.2: Comparison of PSNR

Different defocused regions have slight variations in RMSE and PSNR values. In case of left-right defocused regions, the result is better than that of upper-lower and inner-outer defocused regions. This is because different defocused regions have different clear and blurred pixels due to which the edges are also different in each case. So the gradient operators give slightly different results.

The comparing results show that the proposed image fusion approach has the lower RMSE value and greater PSNR values. It means that the proposed approach is comparatively better than the other methods discussed above.

6.3 Summary

In this chapter analysis of proposed scheme has been done for standard test images (i.e. Lena, Barbara and Peppers) and achieved results compared with previous techniques. The empirical results provide smaller RMSE and higher PSNR values than those provided by some of the previous approaches, which strengthen the idea of using region level image fusion in integration with lucidity decision parameters.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

Nowadays, image fusion has become an emerging and essential tool and shown its power in many fields like image analysis and computer vision, robotics, satellite imagery, night vision applications, remote sensing and medical diagnosis. The goal of image fusion is to combine information from multiple images of the same scene. The result of image fusion is a single image which is more suitable for human perception or further image processing tasks. Generally there are three demands from an image fusion algorithm:

- It should not discard any salient information contained in 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 of imperfections such as noise or mis-registrations.

Image fusion algorithms take a set of input images of the same scene, from different sources, with the aim to obtain new or more precise knowledge about the scene, which is more suitable for human and machine perception or further image-processing tasks such as segmentation, feature extraction and object recognition.

In the past, multifocus image fusion has been carried out using a variety of techniques. All fusion techniques are based on the principle of extracting the valuable information from source images to create fused image containing all objects 'in focus'. The three major techniques are based on region selection methods, multi-scale decomposition methods and learning based methods. In the region selection methods, the input images are primarily separated in regions or into segments using a segmentation technique. From sets of such regions, one region per set is selected based on a sharpness criterion to form the final fused image. The value of the sharpness criterion increases and decreases as objects come into focus and go out of focus, or if the contrast changes in the scene.

In multi-scale decomposition based methods, a variety of fusion schemes has been used. Most of these involve pixel exploitation or substitution at a detail level and the result is changes in the intensity values of the fused image. Learning based methods use training engines, which learn to distinguish between sharp and blurred areas on the basis of the given focused and unfocussed training data sets. We proposed a technique for image fusion that is actually an integration of multi-scale wavelet transform, gradient and mathematical morphology schemes.

As compared with other methods, it is clear that the proposed method outshines other conventional techniques. It uses different block sizes at different levels. The large block size is helpful when image blocks are farther away from boundary between clear and blurred portions. The boundary blocks are clearly distinguished with small block size. In this way, image fusion results are far better than other conventional fusion techniques.

7.2 Future Work

This work presents a technique for image fusion that is actually based on feature level fusion. Salient decision features are extracted from input multi-focused images and on the basis of these features, decision is made about the health of blocks of images. Experimental results show that this approach outshines many other image fusion techniques. Future work to further improve the results will be to embed some intelligent technique to fuse blocks present on the boundary of clear and blurred parts. Since these blocks are fused by conventional wavelet method, even better results can be achieved if this part can be replaced by some intelligent technique like using artificial neural networks or something else.

REFERENCES

1. G. Piella, “Adaptive Wavelets and their Applications to Image Fusion and Compression,” Amsterdam: CWI & University of Amsterdam, 2003.
2. L. Gang and L. Xue-Qin, “Pixel-Level image Fusion Based on Fuzzy Theory,” in International Conference on Machine Learning and Cybernetics, Vol. 3, 2007, pp. 1510 – 1514.
3. H. Chu, J. Li and W Zhu, “A Novel Support Vector Machine-Based Multifocus Image Fusion Algorithm,” in Proceedings of the International Conference on Communications, Circuits and Systems, Vol. 1, 2006, pp. 500 – 504.
4. L. Shutao, T. K. James, and W. Yaonan, “Multifocus image fusion using artificial neural networks,” in Proceedings of the International Conference on Machine Learning and Cybernetics, 2005, pp. 985 – 997.
5. Y. Xia and M. S. Kamel, “Novel Cooperative Neural Fusion Algorithms for Image Restoration and Image Fusion,” in IEEE Transactions on Image Processing, Vol. 16, Issue 2, Feb 2007, pp. 367 – 381.
6. Wen Cao, Bicheng Li, Yong Zhang, “A remote sensing image fusion method based on PCA transform and wavelet packet transform”, Proceedings of the 2003 InternationalConference on Neural Networks and Signal Processing, Beijing, Dec.2003, vol.2, pp.976-981.
7. Zhenhua Li, Zhongliang Jing, Gang Li, Shaoyuan Sun, Henry Leung, “A region-based image fusion using ultiresolution segmentation”, IEEE International Conference on Intelligent Transportation Systems, 2003, pp.96-101.

8. L. Shutao, W. Yaonan, Z. Changfan, "Feature of Human Vision System Based Multi-focus Image Fusion" ACTA Electronica Sininca, Vol. 29, No. 12, pp. 1699-1701, December 2001

9. C. Hua-Wen and L. Shu-Duo, "Image fusion based on addition of wavelet coefficients," in International Conference on Wavelet Analysis and Pattern Recognition, Vol. 4, 2007, pp. 1585 – 1588.

10. Qu Xiaobo and Yan Jingwen, "Image Fusion Algorithm Based on Features Motivated Multi-Channel Pulse Coupled Neural Networks," in the 2nd International Conference on Bioinformatics and Biomedical Engineering, 2008, pp. 2103 – 2106.

11. Xiao-Bo Qu, Guo-Fu Xie, Jing-Wen Yan, Zi-Qian Zhu and Ben-Gang Chen, "Image fusion algorithm based on neighbors and cousins information in nonsubsampled contourlet transform domain," in International Conference on Wavelet Analysis and Pattern Recognition, Vol. 4, 2007, pp. 1797 – 1802.

12. A. Ben Hamza, Yun He, Hamid Karim and Alan Willsky, "A multiscale approach to pixel-level image fusion," in Integrated Computer-Aided Engineering, vol. 12, Issue 2, 2005, pp. 135-146, ISSN: 1069-2509.

13. E. Lallier and M. Farooq, "A real time pixel-level based image fusion via adaptive weight averaging," in Proceedings of the International Conference on Information Fusion, vol. 2, 2000.

14. J. Tian, Y. Zheng and L. Shen, "A Comparative study of Wavelet-based Pixel Level Image Fusion," in Proceedings of the International Conference on Signal and Image Processing, 2004.

15. P. Thevenaz and M. Unser, “A pyramid approach to sub-pixel image fusion based on mutual information,” in Proceedings of the International Conference on Image Processing, vol. 1, 1996, pp. 265-268.
16. Z. Liu, K. Tsukada, K. Hanasaki, Y.K. Ho and Y. P. Dai, “Image fusion by using steerable pyramid,” in Pattern Recognition Letters, vol. 22, Issue 9, July 2001, pp. 929-939.
17. Y. Song, M. Li, Q. Li and L. Sun, “A New Wavelet Based Multi-focus Image Fusion Scheme and Its Application on Optical Microscopy,” in IEEE International Conference on Robotics and Biomimetics, 2006, pp. 401-405.
18. C. Hua-Wen and L. Shu-Duo, “Image fusion based on addition of wavelet coefficients,” in International Conference on Wavelet Analysis and Pattern Recognition, Vol. 4, 2007, pp. 1585 – 1588.
19. Lei Tang and Zong-gui Zhao, “The Wavelet-based Contourlet Transform for Image Fusion,” in the Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, Vol. 2, 2007, pp. 59 – 64.
20. Muwei Jian, Junyu Dong and Yang Zhang, “Image Fusion Based on Wavelet Transform,” in the Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, Vol. 1, 2007, pp. 713 – 718.
21. Gonzalez and Woods, “Digital Image Processing,” 3rd Edition, ISBN: 9780131687288, Prentice Hall, 2008.