To the Graduate Council:

I am submitting herewith a thesis written by Gaurav Ashok Baone entitled "Development of Demosaicking Techniques for Multi-Spectral Imaging Using Mosaic Focal Plane Arrays". I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Electrical Engineering.

Hairong Qi

Major Professor

We have read this thesis
and recommend its acceptance:

Donald W. Bouldin

Seong G. Kong

Accepted for the Council:

Anne Mayhew

Vice Chancellor and Dean of
Graduate Studies

(Original signatures are on file with official student records)

Development of Demosaicking Techniques for Multi-Spectral Imaging Using Mosaic Focal Plane Arrays

A Thesis
Presented for the
Master of Science Degree
The University of Tennessee, Knoxville

Gaurav Ashok Baone
August 2005

Copyright © 2005 by Gaurav Ashok Baone.
All rights reserved.

To my parents, Ashok and Suhasini
for their blessings.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Hairong Qi, for her unwavering support over the years. The association with Dr. Qi has been inspiring and a great learning experience. If not for her, this thesis work and the work behind it would have just remained an unfulfilled dream.
I would like to take this opportunity to thank Dr. Bouldin and Dr. Kong for serving on my thesis committee and for providing me valuable suggestions on my thesis work.
I thank the funding agency, *US Army Space and Missile Defense Command* for sponsoring this project.
I would like to give my special thanks to my labmates at AICIP lab. Hongtao, Lidan, Bala, Sankar, Yingyue, Yang, Woye, Xiaoling and Aruna - thank you for all your support. I am specially grateful to Hongtao for his valuable lectures on China and Chinese food.
I am extremely grateful to my cousin Mugdha and her husband Kaustubh Deshpande for their huge and constant support since my day one in USA.
All my friends at UT, *hats off* to you guys for tolerating me for two years. To name a few of them - Hari, Mardav, Prem, Ravi,the whole Golf-Range "Gulti-Gang" and "Manthan"-the Indian student association at UTK. Most special thanks to my roommate, Ikram, who has showed great patience especially during the last few months of my thesis work.
I dedicate this thesis to my parents and my brother (Chaitanya) who have blessed, loved and supported me, come what may. They have been my ideal and I hope I have managed to live up to their expectations.

Abstract

The use of mosaicked array technology in commercial digital cameras has made them smaller, cheaper and mechanically more robust. In a mosaicked sensor, each pixel detector is covered with a wavelength-specific optical filter. Since only one spectral band is sensed per pixel location, there is an absence of information from the rest of the spectral bands. These unmeasured spectral bands are estimated by using information obtained from the neighborhood pixels. This process of estimating the unmeasured spectral band information is called *demosaicking*. The demosaicking process uses interpolation strategies to estimate the missing pixels. Sophisticated interpolation methods have been developed for performing this task in digital color cameras.

In this thesis we propose to evaluate the adaptation of the mosaicked technology for multi-spectral cameras. Existing multi-spectral cameras use traditional methods like imaging spectrometers to capture a multi-spectral image. These methods are very expensive and delicate in nature. The objective of using the mosaicked technology for multi-spectral cameras is to reap the same benefits it offers in the commercial digital color cameras. However, the problem in using the mosaicked technology for multi-spectral images is the huge amount of missing pixels that need to be estimated in order to form the multi-spectral image. The estimation process becomes even more complicated as the number of bands in the multi-spectral image increases. Traditional demosaicking algorithms cannot be used because they have been specifically designed to suit three-band color images.

This thesis focuses on developing new demosaicking algorithms for multi-spectral images. The existing demosaicking algorithms for color images have been extended for multi-spectral images. A new variation of the bilinear interpolation

based strategy has been developed to perform demosaicking. This demosaicking method uses variable neighborhood definitions to interpolate the missing spectral band values at each pixel locations in a multi-spectral image. A novel Maximum a-Posteriori (MAP) based demosaicking method has also been developed. This method treats demosaicking as an image restoration problem. It can derive optimal estimation result that resembles the original image the best. In addition, it can simultaneously perform interpolation of missing spectral bands at pixel locations and also remove noise and degradations in the image.

Extensive experimentation and comparisons have shown that the new demosaicking methods for multi-spectral images developed in this thesis perform better than the traditional interpolation strategies. The outputs from the demosaicking methods have been shown to be better reconstructed estimates of the original images and also have the ability to produce good classification results in applications like target recognition and discrimination.

Contents

1 Introduction **1**

1.1 Objective 3

1.2 Image Sensors 4

1.2.1 CCD Image Sensors 4

1.2.2 CMOS Image Sensors 5

1.3 Detector Geometry 6

1.4 Mosaic Focal Plane Array Technology 8

1.4.1 Mosaicking Process 9

1.4.2 Demosaicking Process 12

1.5 Contribution of Research 26

1.6 Thesis Outline 26

2 MFPA Technology for Multi-Spectral Images **28**

2.1 Mosaicking of Multi-spectral Images 29

2.2 Demosaicking of Multi-Spectral Images 39

2.2.1 Factors Affecting the Demosaicking Process 40

2.2.2 Demosaicking Methods for Multi-Spectral Images 44

2.3 Problems With Interpolation Based Demosaicking Strategies . . . 52

3 Demosaicking Multi-Spectral Images Using MAP **54**

3.1 Demosaicking as an Image Restoration Problem 55

3.1.1 Theory on Image Restoration 59

3.1.2 Proposed Model for Demosaicking 63

3.2 MAP Technique for Restoration 67

3.2.1 Sensor Model 68
3.2.2 Prior Model 69
3.2.3 Solving the MAP Problem 72

4 Experimental Results and Discussions **76**
4.1 Experimental Image Database 76
4.2 Performance Metrics 77
4.2.1 Reconstruction Accuracy 79
4.2.2 Classification Accuracy Using Spatial Information 80
4.2.3 Classification Accuracy Using Spectral Information 81
4.2.4 Calculation of Classification Accuracy 82
4.3 Demosaicking Results For Multi-Spectral Images 84
4.3.1 Bilinear Demosaicking 86
4.3.2 Cok's Demosaicking 88
4.3.3 Median Based Demosaicking 89
4.3.4 Modified Bilinear Demosaicking 91
4.3.5 Modified Median Based Demosaicking 94
4.4 MAP Based Demosaicking Results 96

5 Conclusions and Future Work **103**
5.1 Contributions of Thesis 103
5.2 Future Work 104

Bibliography **106**

Vita **110**

List of Tables

1.1 Root Mean Square Errors for Color Demosaicking Methods 20

3.1 Comparison Between Image Restoration and Demosaicking Processes 65

4.1 RMSE Values for Different Demosaicking Techniques 95
4.2 Comparison of Degraded Image and MAP Estimated Image (Gaussian Prior Model) 99
4.3 Comparison of Degraded Image and MAP Estimated Image (Gibbs Prior Model with Laplacian kernel) 101
4.4 Comparison of Degraded Image and MAP Estimated Image (Gibbs Prior Model Using Quadratic Variation) 101

List of Figures

1.1 CCD Readout Mechanism [Cutin, 2004] 5
1.2 KAC-0311 CMOS Image Sensor from Kodak (*Courtesy: Eastman Kodak Company*) 6
1.3 Types of Tessellations 7
1.4 Connectivity Definitions in Square and Hexagonal Tessellations . 8
1.5 Color Filter Arrays [FillFactory, 2005] 10
1.6 Block Diagram of Mosaic Technology 11
1.7 Mosaicking Process 12
1.8 Bilinear Demosaicking for Bayer CFA 14
1.9 Bilinear Demosaicking Results for Color Images 16
1.10 Constant-Hue Demosaicking Process Block 18
1.11 Constant-Hue based Demosaicking Results for Color Images . . . 19
1.12 Median-based Demosaicking Results for Color Images 21
1.13 Luminance Channel Interpolation in Gradient-based Demosaicking 22
1.14 Gradient-based Demosaicking Results for Color Images 23
1.15 First Run in Adaptive Color Plane based Demosaicking 24
1.16 Adaptive Color Plane based Demosaicking Results for Color Images 25

2.1 Characteristics of the Human Visual System and the Bayer CFA . 31
2.2 Comparison of Spectral Consistency in CFAs 32
2.3 Sample Checkerboard Pattern 34
2.4 Generation of a Four-Band Mosaic Filter Array Pattern 35
2.5 Four-Band MSFA 37
2.6 Seven-Band MSFA Generation 38

2.7 Sample Seven-Band MSFA 39
2.8 MSFAs for Different Number of Spectral Bands 41
2.9 Example of Inter-Band Dependence in Images 43
2.10 Different Sized Neighborhoods in the Seven Band MSFA 44
2.11 Bilinear Demosaicking of a Seven Band Multi-spectral Image . . . 45
2.12 Cok's Demosaicking of a Seven Band Multi-spectral Image 47
2.13 Median Based Demosaicking for Seven Band Multi-spectral Images 48
2.14 Modified Bilinear Demosaicking Process 50

3.1 MFPA Block Diagram 56
3.2 General Block Diagram of an Image Restoration Problem 56
3.3 Demosaicking In Presence of Degradation 58
3.4 Demosaicking as Image Restoration Process 64
3.5 Inverted Gaussian as Penalty Function 71

4.1 Database of Multi-Spectral Images 78
4.2 Block Diagram for Invariant Moments Calculation 82
4.3 Calculation of Classification Accuracy 83
4.4 Experimental Process 85
4.5 Seven Bands of the Original Multi-Spectral Image 85
4.6 The Mosaicked Image 85
4.7 Bilinear Demosaicking Results 86
4.8 Classification Accuracy Curves for Bilinear Demosaicked Output . 87
4.9 Zipper Effect in Bilinear Demosaicked Output 88
4.10 Cok's Demosaicking Results 89
4.11 Classification Accuracy Curves For Cok's Demosaicked Output . . 90
4.12 Median Based Demosaicking Results 91
4.13 Variation of Edge Information in the Multi-Spectral Image 91
4.14 Classification Curves for Median Based Demosaicked Output . . . 92
4.15 Modified Bilinear Demosaicking Results 93
4.16 Classification Curves for Modified Bilinear Demosaicked Output . 93
4.17 Modified Median Based Demosaicking Results 94

4.18 Classification Curves for Modified Median Based Demosaicked Output 95
4.19 MAP Experimentation Process 97
4.20 Original Image has been Degraded with Gaussian Noise and Blur. The Degraded Image is then Mosaicked to form the Degraded Mosaicked Image. 98
4.21 MAP Estimate Results Using Gaussian Prior Model 99
4.22 MAP Estimate Results Using Gibbs Prior Model (Laplacian Kernel) 100
4.23 MAP Estimate Results Using Gaussian Prior Model (Quadratic Variation) 101

Chapter 1

Introduction

In every phenomenon the beginning remains always the most notable moment.

– Thomas Carlyle

In the past two decades, most of the major technological breakthroughs have been made towards making the fundamental shift from analog technology to the digital technology. This technological shift has changed the way in which the world handles visual and audio information. The digital camera is one of the most remarkable examples of the equipments that have made this shift a great success. The ability of this equipment to convert analog electromagnetic signals to a viewable digital representation, has made the digital cameras more favorable than the traditional image acquisition systems. The use of digital cameras totally eliminates the laborious mechanical and chemical processes involved in creating a picture using traditional cameras. A scene can be instantly captured, viewed and, post-processed using a computer. In short, the use of digital cameras has revolutionized the area of photography and made it even more convenient and reliable.

The credit for the popularity and versatility of digital cameras goes to the technology that is used inside the camera. The use of sensors that can convert light into electrical charges has made it possible to replace the traditional camera

technology. The image sensor employed by most of the digital cameras is a Charge Coupled Device (CCD). Some of the cameras also make use of sensors that use the Complementary Metal Oxide Semiconductor (CMOS) technology. These sensors are arranged in the form of an array inside the camera. The image formation process inside the camera involves light-sensing at each location of the image by the image sensors and then combining their outputs in a proper fashion to form the output image that can be perceived by the human eye. In case of color cameras, each pixel in the image has to have information from all the three primary visual bands (Red, Green and Blue). To tackle this problem, each location in the imaging array inside the camera is made to have three image sensors each sensitive to the three primary visual spectral bands. This enables each pixel to register information from the red, green and blue spectral bands from the incoming light. However, the problem with using such a technique for image formation is that, as the demand for greater image resolution increases, the number of pixels in the image also become higher which results in a greater number of locations of the imaging array. This in turn increases the total number of imaging sensors required for forming the color image. This affects the cost and the size of the image capturing equipment. Due to the presence of a huge number of sensors, the pixel registration also tends to become a little less efficient. This problem was tackled by the use of a new technology called the Mosaic Focal Plane Array (MFPA) technology. The MFPA technology enables the use of just one image sensor per pixel location. The technology offers to reduce the size and cost of the equipment and at the same time offering an almost similar image resolution and quality. MFPA also eliminates the pixel registration problem that occurs in the traditional digital camera technology. Today's commercial digital camera industry heavily relies on the MFPA technology. The MFPA technology has brought about a technological change that is commercially more viable and cost effective.

The applications of digital images have surpassed the boundaries of visible electromagnetic spectrum. Today, digital images are used to analyze entities that cannot be seen by a human eye. For example, the use of infrared spectral band to distinguish crop types in agriculture. The crop types cannot be distinguished using the traditional color images. This creates a need for an image that can

accommodate the information from the infrared spectral band too. This is generally achieved by forming an image with information from all the required spectral bands. Such an image which consists of information from more than the three traditional visible spectral bands is called a **Multi-Spectral** image. Multi-Spectral images have wide range of applications ranging from agriculture, medicine, defense, etc. The use of multi-spectral images has made night-vision possible for defense related equipments. In agriculture, the use of multi-spectral images has made analysis of crop sustainability, irrigation planning and storm damage assessment easier. However, the technology for multi-spectral image acquisition still is in its primitive form. The existing multi-spectral cameras use traditional methods to capture and form the multi-spectral images. The multi-spectral cameras lack the benefits offered by the popular MFPA technology that is used for digital color cameras.

1.1 Objective

In this thesis we would explore the possibility of using the MFPA technology for building a multi-spectral camera. The purpose of adapting the MFPA technology as a possible new technology for multi-spectral cameras is to make the systems cheaper, smaller, and mechanically more robust than any existing multi-spectral image acquisition alternative. Thereby, this technology can provide the benefits to multi-spectral cameras which are now being reaped by the digital color camera industry. The thesis will involve the developments of new interpolation strategies which can be used to recreate the multi-spectral image from the registered pixel values in the imaging array from various spectral bands. Possible extensions of existing demosaicking algorithms have been explored and implemented. Another aspect of this work is to improve the quality of the multi-spectral images in the presence of degradations and noise which tend to creep in during the image acquisition process. This research work has been funded by the *US Army Space and Missile Defense Command* and is a collaborative effort of University of Tennessee, Knoxville and North Carolina State University.

1.2 Image Sensors

The key difference between a digital camera and a film camera is that the digital camera has image sensors instead of a film to capture the scene. The image sensors convert light to electrical charges. Popular image sensor technologies include the **Charge-Coupled Device** (CCD) sensors, the **Complementary Metal Oxide Semiconductor** (CMOS) devices, the **Charge Injection Device** (CID) and the **Amorphous Silicon** (a-Si) Sensor. The image sensors provide various advantages compared to the traditional film photograph. Firstly, they are faster than film as they are capable of generating a digital image almost instantaneously compared to the laborious process that has to be followed in case of a film. Secondly, image sensors offer higher sensitivity. The *quantum efficiency*, or the ability to record the incoming light, is about 90% compared to a quantum efficiency of 0.5% in traditional photography [Qi, 1999]. Lastly, image sensors are linear in nature. This linearity makes the digital image acquisition process more stable and greatly reliable compared to the non-linear traditional photography methods. In this section, we will give a brief overview of the two most popular image sensor technologies.

1.2.1 CCD Image Sensors

A CCD is a silicon wager which is capable of converting incoming photons to electrons which could be stored and counted. The CCD was invented by George Smith and Willard Boyle at Bell Labs in 1969 [Cutin, 2004]. The CCD is an array of metal-oxide-semiconductor (MOS) capacitors which can accumulate and store charge due to their capacitance [Audley, 1997]. The CCDs get their name from the way the charge is read after an exposure. The array is read out by transferring the charge from one MOS capacitor to its neighbor on one side. Fig. 1.1 illustrates the CCD operation. The charges on the first row are transferred to a read-out register. Then they are fed to the amplifier unit which provides the required amplification. They are then converted to digital sequences by the analog-to-digital converter. This process is followed for each row in the array. After the completion of each row, the charges in the read-out register are removed and replaced by the charges

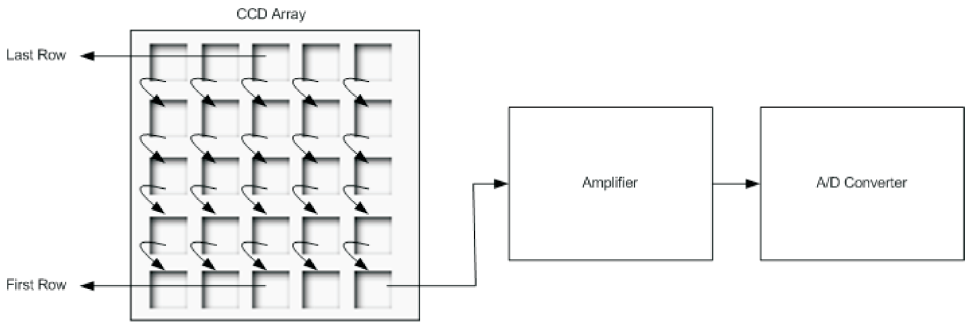


Figure 1.1: CCD Readout Mechanism [Cutin, 2004]

from the next row. In this way the charges on each row are *coupled* to those on the row above so when one row moves down, the next one moves down to fill its old space in the read-out register. In this way each row is read one at a time. The CCD technology remains the dominant technology in the image sensor market. According to a survey [Qi, 1999], the CCD sensors have a 79% of market share.

1.2.2 CMOS Image Sensors

The main problem with the CCD sensors is that they are not cost-effective. They have been produced in foundries using specialized and expensive processes that are specific only to make CCDs. Meanwhile, in the case of CMOS image sensors, there are more and larger industries which have already been making products that use the CMOS technology to make chips for computer processors and memory. There are two kinds of CMOS image sensors - *Passive Pixel Sensors* and *Active-Pixel Sensors* [Cutin, 2004]. In passive-pixel CMOS sensors, a photosite converts photons into an electrical charge. This charge is then carried off the sensor and amplified. These sensors are small in size. However, the main problem with these kind of sensors is the noise that occurs in the image acquisition process. The active-pixel sensors reduce the noise associated with the passive-pixel sensors. Each pixel has a special in-built circuitry that determines the noise and cancels it out. The performance of this technology is comparable to the CCD image sensors and also allows for larger image array and higher resolution. Fig. 1.2 shows the Kodak CMOS KAC-0311 image sensor. The sensor can accommodate 640×480

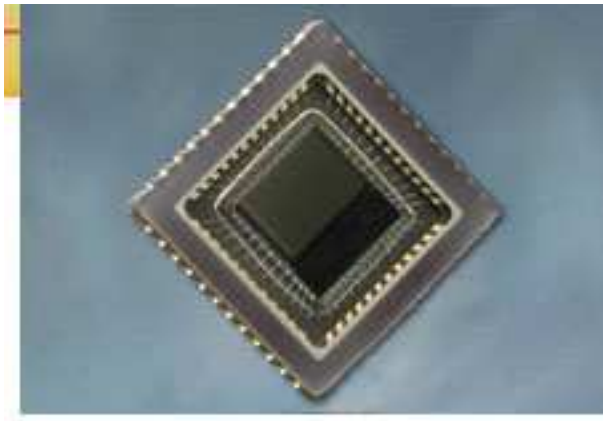
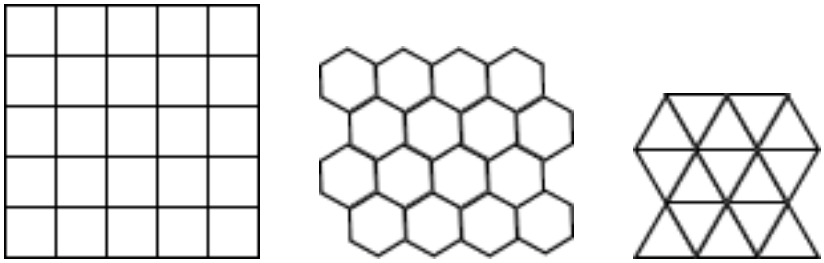


Figure 1.2: KAC-0311 CMOS Image Sensor from Kodak (*Courtesy: Eastman Kodak Company*)

active pixel elements [Company, 2004].

1.3 Detector Geometry

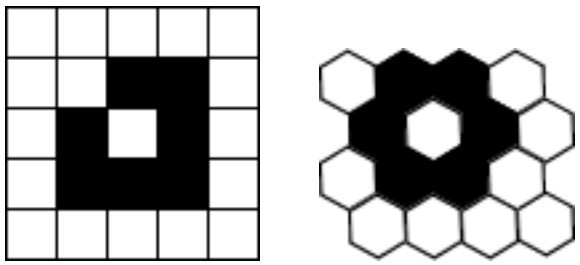
The detector geometry plays a very important role in the image acquisition system. In general, there are only three possible ways in which a plane can be tiled with regular shapes. These are illustrated in Fig. 1.3. The most popular pixel geometry that is used in imaging systems is the square tessellations. One of the main reasons for the choice of square-shaped tessellation for pixels is that the pixels are addressed based on the orthogonal cartesian coordinate system which is the most widely used coordinate system. It has also been debated that square tessellations are the best model for the human visual system because the human vision is supposed to observe straight vertical features. However, the hexagonal tessellations have also been explored and proved to be better than the square tessellations. The triangular tessellations is considered to be the dual of the hexagonal tessellations. The hexagonal tessellations have some advantages over the square tessellations. *Firstly*, the hexagonal system is supposed to model the human visual system the best. This is because the transformation from image to cortical representation consists of two stages. The first stage is between the image and the retinal neural image and the second between the retinal neural



(a) Square Tessellation (b) Hexagonal Tessellation (c) Triangular Tessellation

Figure 1.3: Types of Tessellations

image and the cortical neural image [Wensel et al., 1990]. The distribution of the ganglion cells is rotationally invariant about the center of the visual field. That is, the cells on a circle of given radius are uniformly distributed. Each unit has a center-surround structure referred to as a *heptuplet*, consisting of itself and the six neighboring units whose centers form a hexagon. This is the reason the receptors in the human visual system are arranged in hexagonal tessellation. Thus, the hexagonal tessellation best maps human visual system. *Secondly*, it has been proved that if a two-dimensional signal is sampled using hexagonal sampling then it takes about 13.4% lesser samples than the rectangular sampling [Dudgeon and Mersereau, 1984]. *Thirdly*, there is no connectivity ambiguity in the hexagonal tessellation. The square tessellation is said to have two types have connectivities - the edge-to-edge connectivity and the corner-to-corner connectivity [Rosenfled, 1970]. These two definitions of connectivity in square tessellations leads to the connectivity ambiguity. The figure 1.4(a) displays the connectivity ambiguity. We see that even though the inner square pixel is in a closed path according to the edge-to-edge connectivity definition still the background is connected to the foreground through the open corner.This shows that the foreground and background are connected and the inner pixel is still not in a closed loop. But this type of ambiguity does not arise in the case of hexagonal tessellation. The reason for that is because there is only one definition for connectivity which is related to the 6-neighborhood (see Fig. 1.4(b)). *Next*, the hexagonal tessellation has been proved to be the most optimal for thinning algorithms and edge detection. *Finally*, the advantage of equidistant neighbors in hexagonal tessellations



(a) Connectivity Ambiguity In Square Tessellations

(b) Perfect Connectivity In Hexagonal Tessellation

Figure 1.4: Connectivity Definitions in Square and Hexagonal Tessellations

proves beneficial in the interpolation process. The interpolation process takes the neighborhood pixel intensities into consideration for estimation of the missing pixel value. Now, in a square tessellation when an eight-neighborhood definition is considered the neighbors at the diagonals are farther (distance wise) than the vertical and the horizontal ones. Thus, the characteristics of the diagonal neighbors are less similar to the center pixel than to the rest of the neighbors. This problem does not occur in hexagonal tessellation because in hexagonal tessellation all the six neighbors are at the same distance and thus while performing interpolation, each of the neighborhood pixel contributes equally towards the missing pixel intensity calculation.

1.4 Mosaic Focal Plane Array Technology

Mosaic focal plane array technology has become mainstay in the making of commercial digital cameras. This has been possible because the technology offers many advantages compared to the traditional technologies that have been used in digital cameras. Some of the advantages include lesser equipment costs, greater robustness and better pixel registration. These benefits will become more evident as we go into details of this technology.

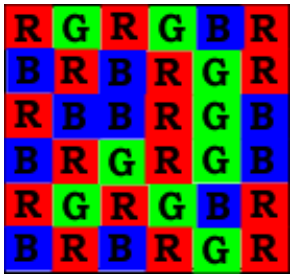
The Mosaic Focal Plane Array technology uses mosaicked sensors to capture images. A mosaicked sensor is a monolithic array of many sensors, arranged in a particular geometric pattern. This pattern is generally called the mosaic filter

array. Some of the popular filter arrays for color-image cameras have been shown in Fig. 1.5. These filter arrays are called color filter arrays (CFA). Sensors for specific wavelengths of the spectral band are arranged in the photo-sites according to the CFA pattern. Fig. 1.6 gives a block diagram of the mosaic technology.

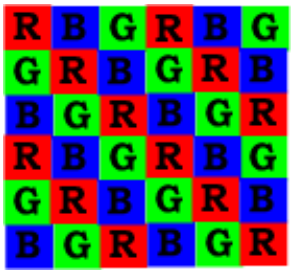
The block diagram gives us an overall picture of the image acquisition process that takes place in a camera. The actual scene is captured using the mosaicked sensors. This process of capturing the actual scene onto a set of wavelength-specific sensors is called Mosaicking. The image capturing process is such that only one specific spectral band is captured per pixel location. This helps in reducing the total number of sensors required for the image acquisition process, thus reducing the size and cost of the equipment. The output of the mosaicking process produces an image which is a collection of intensity values produced by the various sensors on the mosaicking array. This mosaicked image is then given as input to the demosaicking block. The demosaicking block tries to estimate a multi-channel image from the mosaicked image. Since, each pixel in a mosaicked image has information of only one particular spectral band, the demosaicking process uses neighborhood pixels to estimate the prospective intensity values of the missing bands at each pixel location. This process is called interpolation. There are many interpolation algorithms that have been developed and are being widely used in commercial digital cameras. In the following sub-sections we will look into more details of the mosaicking and demosaicking processes. The discussion will mainly focus on the use of mosaic focal plane array technology in digital cameras for color images.

1.4.1 Mosaicking Process

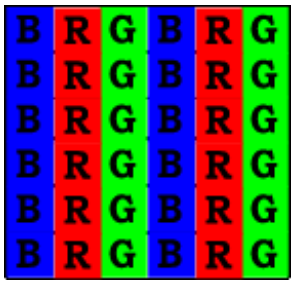
Mosaicking is a technique that is used to capture a multi-channel or a multi-band image using only one sensor per pixel [Ramanath et al., 2004]. To form a multi-band image, each pixel in the image needs to have information from all the spectral bands. For example, if a 3-band image has to be formed, then each pixel in the 3-band image must contain information pertaining to all the three spectral bands. In case of mosaicking, the required spectral bands are subsampled such that the



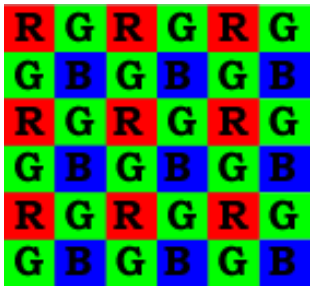
(a) Pseudo-randomly generated CFA



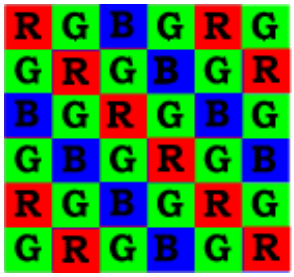
(b) Diagonal Stripe CFA



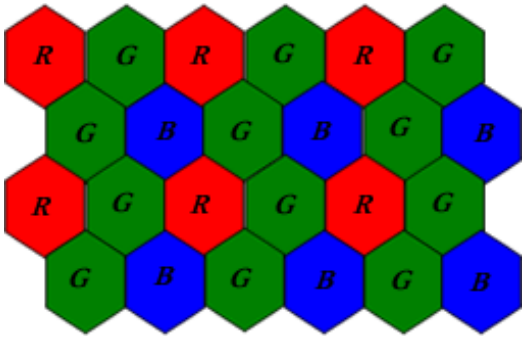
(c) Vertical Stripe CFA



(d) Bayer CFA



(e) Diagonal Bayer CFA



(f) Bayer CFA (in hexagonal tessellations)

Figure 1.5: Color Filter Arrays [FillFactory, 2005]

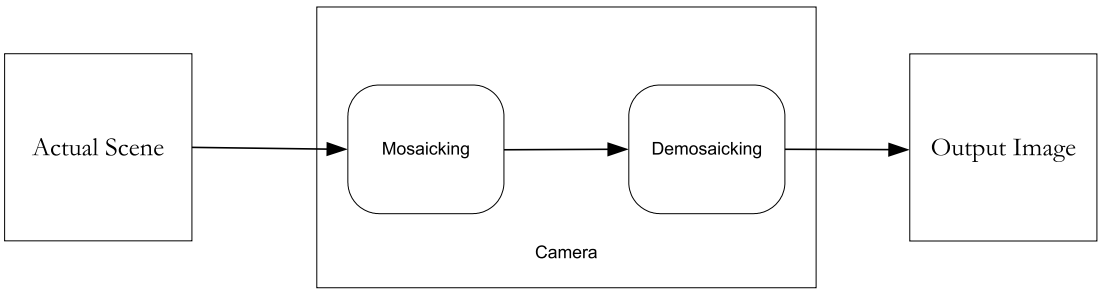


Figure 1.6: Block Diagram of Mosaic Technology

image can be formed by using only one sensor per pixel instead of using multiple sensors at each of the pixel location. This sampling is achieved by overlaying a filter array on top of the sensor substrate. Ideally, the image formation process requires each pixel to have information from all the spectral bands. This is made possible by placing all the required photo-sensors sensitive to all the spectral bands at each pixel location. The sensors would then sense the incoming light and record the intensity values at each sensor location. The problem with this technique is that the number of sensors required for the formation of the whole image is very huge. Thus, increasing the cost of the equipment to a large extent. Another major problem is the one of pixel registration. Since, each pixel location has a cluster of sensors packed in a small space, the probability of sensing the exact intensity of each of the spectral bands is reduced to a great extent. This consequently degrades the quality of the image formed. The construction of this kind of image acquisition equipment makes it mechanically less robust. Any disturbance to the photo-sensors, due to external force, can result in erroneous sensing by the sensors. These drawbacks have been addressed by the mosaicking process.

As discussed earlier, the mosaicking process uses a mosaic filter array to sense one spectral band per pixel location. Fig. 1.7 explains the mosaicking process for a three band color image. The actual scene is captured by the CFA in the form of a mosaicked image. The mosaicked image is a collection of all the spectral bands in a specific pattern which is dictated by the pattern of the color filter array. The heart of the mosaicking process lies in the design of the color filter array.

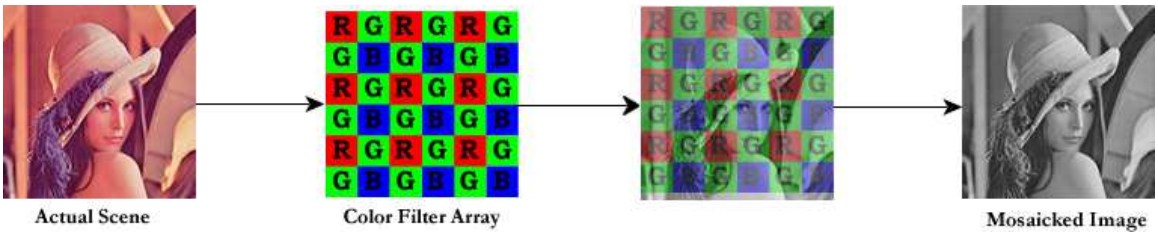


Figure 1.7: Mosaicking Process

Color Filter Array The color filter array (CFA) forms the most important part of the mosaicking process. The CFA enables sensing just one spectral band per pixel location instead of sensing all the spectral bands at each of the pixel locations. The ultimate aim of using a CFA in the image acquisition process is to reduce the size and cost of the equipment. At the same time, the CFA must be designed carefully so that it is easier to form an accurate estimation of the scene being captured in the form of a digital color image. This makes the designing process of a CFA a serious matter and thus needs consideration of various factors like frequency or probability of appearance of spectral bands, spectral consistency, uniform spectral distribution and the most important, the CFA must provide means for easy interpolation to form the final multi-band image [Miao et al., 2003]. We will look into the details of each of these factors in Sec. 2.1.

1.4.2 Demosaicking Process

Demosaicking[1] is the reverse process of Mosaicking. The main aim of demosaicking is to estimate the color image (multi-band image) from the mosaicked image formed by the mosaicking process. The mosaicked image contains information about all the spectral bands distributed throughout the image in a specific pattern which is defined by the CFA. Now, to form the whole image from the dis-

[1]It is customary in research literature to mention the ambiguity in spelling of the word *demosaicking*. Some authors prefer spelling the word as *demosaicing*. However, since the issue still remains unresolved, the manner in which the word is spelt has been left to the author's discretion.

tributed spectral band information we need to estimate the missing band values at each of the pixel location such that the ouput is as close as possible to the actual scene. Traditionally, interpolation techniques have been used for demosaicking color images. These techniques use neighborhood information to estimate the missing spectral band information at each of the pixel locations. In this section, we present the popular interpolation techniques [Ramanath et al., 2002] that have been used to demosaic images mosaicked by the Bayer CFA.

Ideal Interpolation The concept of ideal interpolation is only of theoretical importance as it cannot be implemented practically. This is shown with the help of mathematical modelling. For the ease of interpretation, we only consider the one-dimensional case [Roberts, 2003]. Before going into the mathematical details of one-dimensional signals, here is a description of the interpolation process in case of images. The actual scene is a two-dimensional continuous signal, $f(x, y)$, which is sub-sampled during the mosaicking process using the CFA. It has to be restored back into a high-resolution digital image by interpolating the missing samples in the mosaicked image.
A Continuous Time (CT) signal $x(t)$ is sampled with a sampling rate of f_s.The impulse-sampled signal is mathematically denoted as in Eq. (1.1)

$$x_\delta(t) = \sum_{n=-\infty}^{\infty} x(nT_s)\delta(t - nT_s) \tag{1.1}$$

where $T_s = \frac{1}{f_s}$. The reconstruction of the original signal is performed using an ideal lowpass filter. In the frequency domain, the lowpass filter cuts off above f_m (f_m being the highest frequency present in the original signal) and below $f_s - f_m$. Then

$$X(f) = T_s rect(\frac{f}{2f_c})X_\delta(f) \tag{1.2}$$

where f_c is the cutoff frequency. $f_m < f_c < (f_s - f_m)$. The equivalent in the time domain will be

$$x(t) = 2\frac{f_c}{f_s} sinc(2f_c t) * x_\delta(t) \tag{1.3}$$

So we see, in the time domain the interpolation process uses the convolution operation which is nothing but the shifted sum of the signals involved over infinite time space. That is, to obtain perfect signal reconstruction, we need to perform the process over all the infinite samples of the sampled signal. This is not possible practically as we only have limited number of samples. Thus, the implementation of the ideal reconstruction filter is limited only to theoretical reasoning of the interpolation process.

Bilinear Demosaicking Bilinear demosaicking is considered to be the most basic and the most simple way to *demosaic* a given image [Longere et al., 2002]. This method of demosaicking depends only on intensity values of the same spectral band, i.e, it does not take into consideration the possible correlation between spectral bands. The interpolation of a missing spectral band value is performed by considering the pixel values of the same spectral band in the neighborhood of the pixel location. The bilinear demosaicking algorithm is illustrated in Fig. 1.8.

Take, the pixel location R_{53}, as an example. This pixel location has recorded the intensity value of only the red spectral band. The missing spectral bands, green and blue are estimated by considering a 3×3 neighborhood and averaging the pixel values with the corresponding spectral band. Thus, the green and blue

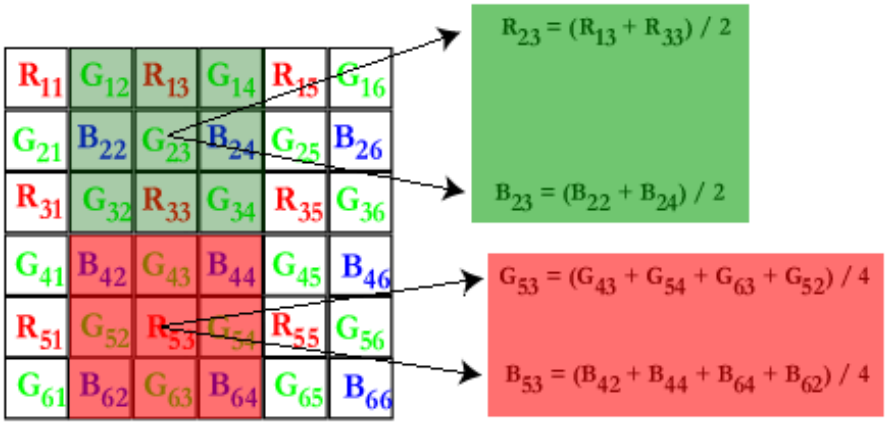


Figure 1.8: Bilinear Demosaicking for Bayer CFA

bands' values will be

$$G_{53} = \frac{1}{4}(G_{43} + G_{54} + G_{63} + G_{52})$$

$$B_{53} = \frac{1}{4}(B_{42} + B_{44} + B_{64} + B_{62})$$

The bilinear demosaicked output can be observed in Fig. 1.9(c). The bilinear demosaicking process introduces step-like artifacts at the edge locations in the image referred to as the *zipper effect*. The artifacts can be closely observed in Fig. 1.9(d). The remedy to the *zipper effect* problem is to consider inter-band correlation in the demosaicking process.

Constant-hue based Demosaicking The constant-hue based demosaicking was proposed by Cok and was one of the first few methods used in commercial digital still cameras. The method focuses on reducing color artifacts generated in the process of reconstruction of the demosaicked output from the mosaicked image. Extending the discussion on bilinear demosaicking, the main cause for the occurrence of *zipper effect* is that the bilinear demosaicking algorithm takes all the spectral bands to be independent from each other. However, it has been empirically proved that color images have cross correlation between the red, green and blue bands. This means that if the inter-band relationship is taken into consideration, the output of the bilinear demosaicking method can be improved to a great extent. Cok [Cok, 1987] proposes that the main cause for the generation of artifacts, caused by the interpolation process, are due to abrupt hue changes. If the hue changes are made to occur gradually then the appearance of color fringes in the image can be reduced to great extent.
In general, hue is defined as the quality of a color as determined by its dominant wavelength. In case of Cok's method, hue is defined by a set of $\frac{R}{G}$ and $\frac{B}{G}$ ratios, where R,G,B stand for intensity values corresponding to the red, green and blue spectral bands respectively. The definition has to be modified when $G = 0$. The red and blue bands are considered to be chrominance channels and the green is assigned to the luminance channel. The method is actually a two pass process, the first being the interpolation of the hue values and then finding the interpolated



(a) Original image considered for experimentation



(b) Mosaicked image obtained from Bayer CFA



(c) Bilinear Demosaicked Output



(d) Zoomed version of the demosaicked output showing the *zipper effect*

Figure 1.9: Bilinear Demosaicking Results for Color Images

chrominance values from the already interpolated hue values. Consider, two pixel locations (a, b) and (c, d) in an image with uniform hue, then the relationship between the chrominance and luminance components at the two pixel locations is given as follows:

$$\frac{R_{ab}}{G_{ab}} = \frac{R_{cd}}{G_{cd}}$$

i.e.,

$$\frac{R_{cd}}{R_{ab}} = \frac{G_{cd}}{G_{ab}} \tag{1.4}$$

Now, if the red value at the pixel location (c, d) is unknown then it can be calculated by rearranging Eqn. (1.4) as

$$R_{cd} = G_{cd} \cdot (\frac{R_{ab}}{G_{ab}})$$

where, R_{ab} and G_{ab} are the measured chrominance and luminance values at the pixel location (a, b). But, usually the images do not have constant hue characteristics. In such non-uniform hue images, the neighborhood information is considered to ensure smooth hue changes throughout the interpolation process. Eqn. (1.5) shows the process of finding the unknown chrominance values in a non-uniform hue image.

$$R^i_{cd} = G_{cd} \cdot (\frac{R_{ab}}{G_{ab}})^i \tag{1.5}$$

where the superscript i represents the interpolated hue value between neighboring chrominance sample locations and G_{cd} is the luminance value at the interpolation location. Now, when it comes to the implementation of the algorithm on an image, the process boils down to a two pass process as illustrated in Fig. 1.10. The first pass interpolates the luminance (i.e. green band) values all over the image. Bilinear interpolation is used to interpolate the green band values. This is followed by interpolating the chrominance values by taking the neighborhood pixels into consideration. As we see from the figure, in the first pass all the missing green pixels are estimated using bilinear interpolation. A 3×3 neighborhood is considered to perform the averaging operation. Next, the final image is formed by estimating the missing chrominance values using the interpolated luminance

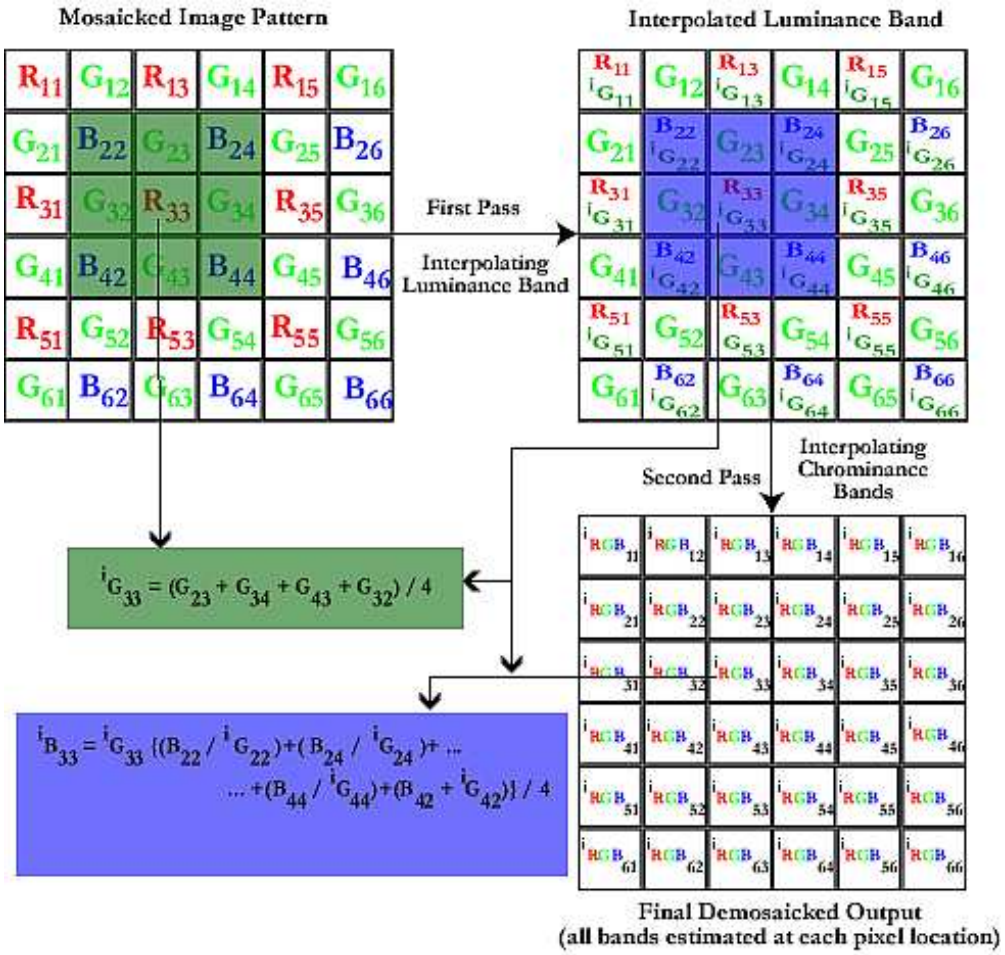


Figure 1.10: Constant-Hue Demosaicking Process Block



(a) Mosaicked image obtained from Bayer CFA



(b) Constant-Hue based Demosaicked Output

Figure 1.11: Constant-Hue based Demosaicking Results for Color Images

information. Fig. 1.11 shows the result obtained after performing the constant-hue demosaicking. Comparing the result with the bilinear demosaicked output from Fig. 1.9, we see that the *zipper effect* in the bilinear demosaicked output has been reduced to a great extent in the case of constant-hue demosaicked image. However, when we compare the original image to the constant-hue demosaicked image, we see that the reconstruction process has not been very successful in creating a perfect visual output. For more clarity on the type of output obtained, refer to Table 1.1. The table lists the root mean square errors (RMSE) between each of the demosaicking outputs and the original image. We observe, from the table that, the constant hue based demosaicking method produces an output that has a RMSE of 23.1798. Ideally, for an output which is an exact replica of the original image, the RMSE value is 0. Comparing, the RMSE values of the Constant-Hue output to the Bilinear demosaicked output, we see that the RMSE of the bilinear demosaicked output is higher than the constant-hue based method. This means, the constant-hue method performs better than the bilinear demosaicking method but still needs a lot of improvement.

Table 1.1: Root Mean Square Errors for Color Demosaicking Methods

Demosaicking Method	**Root Mean Square Error**
Bilinear Demosaicking	23.7870
Constant-Hue Based Demosaicking	23.1798
Median-Based Demosaicking	22.9198
Gradient Based Demosaicking	11.8478
Adaptive Color Plan Based Demosaicking	**10.3228**

Median-based Demosaicking The median-based demosaicking method was invented by Freeman. Freeman [Freeman, 1988] proposes that since the luminance band is the highest sampled band in the Bayer CFA, it is more accurately recreated using bilinear interpolation than the rest of the bands. The higher sampling rate of the luminance band makes the band have more neighbors than the other bands, and thus the probability of estimating the most accurate missing value is greatly increased. To ensure proper estimation of the other bands, Freeman has proposed a two pass algorithm that uses median filtering and produces a better result than the bilinear demosaicking method. Firstly, all the three bands are bilinearly interpolated over the whole image. This output is the same as the bilinear demosaicked output. In the second pass, two difference images $R-G$ and $B-G$ are created, where R, G, B represent the red, green and blue channels of the bilinear demosaicked output. These difference images are then median filtered. Median filtering is a method in which each pixel is assigned to the median of its neighborhood pixels' value. It is a blurring technique which is popularly used in image processing applications. In this context, the median filtering brings about a continuity to the intensity values in the $R-G$ and $B-G$ difference images. Median filtering should be applied only at locations where there are missing R and B values. Next, the interpolated green band image is added to the difference images and thus we get the estimated red and blue channels. Fig. 1.12 shows the result obtained after median-based demosaicking. A 5×5 median filter was used to obtain this result. Referring to Table 1.1, we see that the RMSE of the median based output is 22.9198. There has been an improvement in the result compared to the constant-hue based demosaicking output, but the method still needs to be improved to produce a better result. The drawback of this method is that, the



(a) Mosaicked image obtained from Bayer CFA

(b) Median-based Demosaicked Output

Figure 1.12: Median-based Demosaicking Results for Color Images

method works best only for color images which have similar intensity mean values for all the three bands. Thus, before applying the demosaicking method one has to scale all the bands in the image such that all the bands have similar mean values.

Gradient-based Demosaicking The gradient based demosaicking method was invented by Laroche and Prescott. This method is being actually used in Kodak DCS 200 Digital Camera System. This demosaicking algorithm focuses on improving the output of the bilinear demosaicking output by using the edge information in an image [Laroche and Prescott, 1994]. The method takes advantage of the fact that the human eye is most sensitive to luminance changes. Unlike, bilinear demosaicking, the interpolation is performed by taking the edge information of the luminance channel into consideration. This helps in reducing the edge artifacts which are one of the byproducts of the bilinear interpolation process. The algorithm is a two pass process. The first pass, involves interpolation of the green channel. From Fig. 1.13, we see that at location $(3, 3)$ the missing green band value is calculated by using two gradient values. The hg and vg are two gradient values corresponding to the horizontal and vertical edges. Now, depending

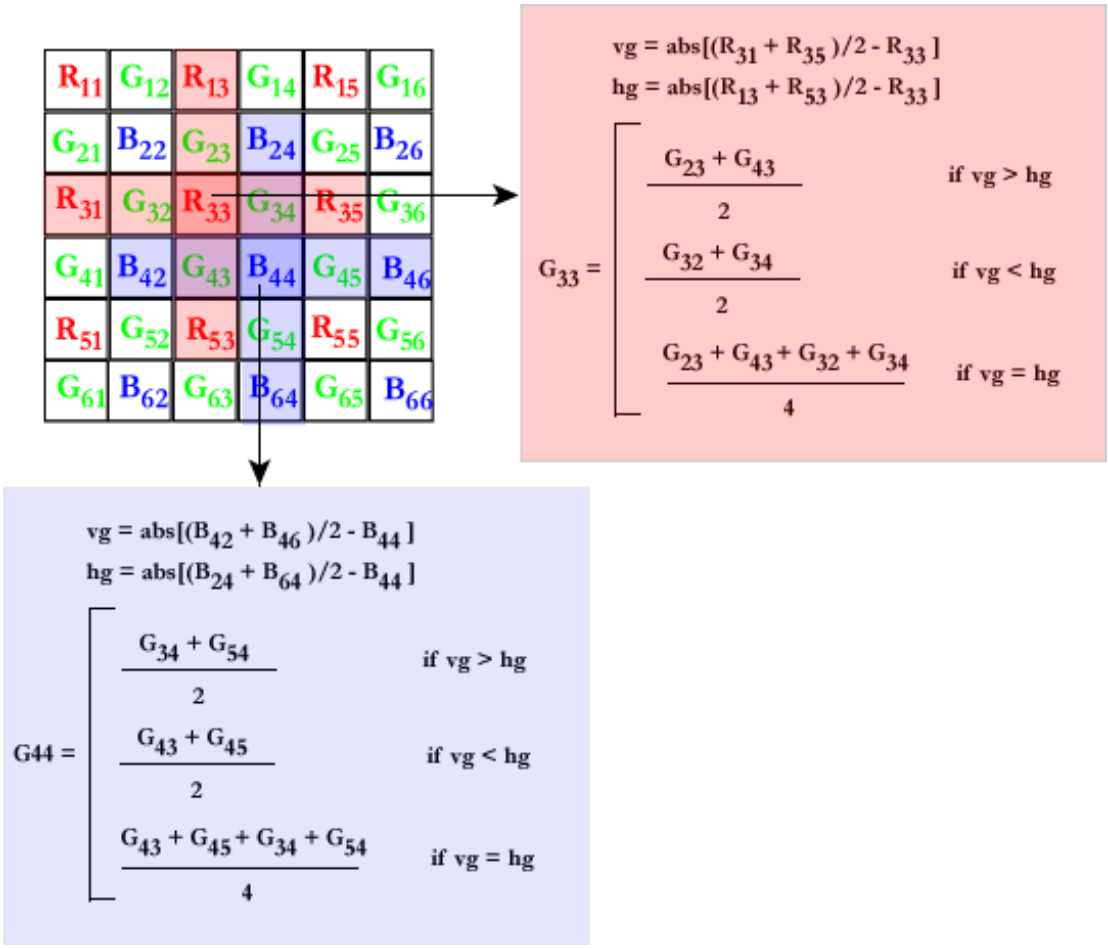


Figure 1.13: Luminance Channel Interpolation in Gradient-based Demosaicking



(a) Mosaicked image obtained from Bayer CFA

(b) Gradient-based Demosaicked Output

Figure 1.14: Gradient-based Demosaicking Results for Color Images

on the strength of the gradients the interpolation is performed in that particular direction. This helps in restoring the existing edge information of the image. The second pass, involves interpolation of the chrominance bands. This process uses the difference images $R - G$ and $B - G$ to interpolate the chrominance values. For eg.,

$$B_{34} = \frac{(B_{24} - G^i_{24}) + (B_{44} - G^i_{44})}{2} + G^i_{34}$$

$$B_{43} = \frac{(B_{42} - G^i_{42}) + (B_{44} - G^i_{44})}{2} + G^i_{43}$$

$$B_{33} = \frac{(B_{22} - G^i_{22}) + (B_{24} - G^i_{24}) + (B_{42} - G^i_{42}) + (B_{44} - G^i_{44})}{4} + G^i_{33}$$

where the superscript i denotes the already interpolated luminance values. The result obtained for the gradient-based demosaicking process is shown in Fig. 1.14. We can observe that the quality of the output has improved to a great extent compared to the previous demosaicking methods. This improvement is clearly seen in the RMSE value of the output (*see Table 1.1*). The RMSE value has greatly reduced compared to the one corresponding to the median based demosaicking method.

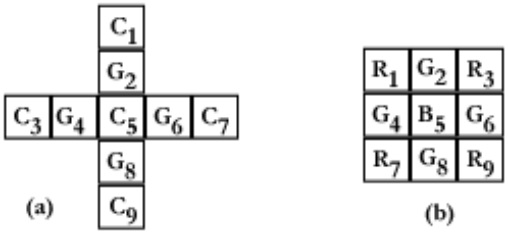


Figure 1.15: First Run in Adaptive Color Plane based Demosaicking

Adaptive Color Plane Demosaicking The adaptive color plane demosaicking technique was developed by Hamilton and Adams [Hamilton and Adams, 1997]. This method employs classifiers similar to the gradient-based demosaicking method but modified to accommodate first and second order derivatives. This process has three runs. The first run interpolates the luminance channel and the next two runs estimate the missing two chrominance band pixels. Consider a Bayer CFA neighborhood in Fig. 1.15 (a) G_i is a green pixel and C_i is a chrominance pixel (i.e, Red or Blue pixel) of the same type. The classifiers are given by

$$hg = abs(-C_3 + 2C_5 - C_7) + abs(G_4 - G6)$$

and

$$vg = abs(-C_1 + 2C_5 - C_9) + abs(G_2 - G8)$$

We observe that the classifiers contain second order derivative terms for chromaticity data and gradients of the luminance data. Now, the first run of interpolating the luminance band is done depending on the orientation of the edge at that pixel location. For example, in our case, G_5 will be determined as given by the following. $G_5 = \left\{ \begin{array}{ll} \frac{G_4+G_6}{2} + \frac{-C_3+2C_5-C_7}{2} & \text{if hg} < \text{vg} \\ \frac{G_2+G_8}{2} + \frac{-C_1+2C_5-C_9}{2} & \text{if hg} > \text{vg} \\ \frac{G_2+G_8+G_4+G_6}{4} + \frac{-C_1-C_9+4C_5-C_3-C_7}{4} & \text{if hg} = \text{vg} \end{array} \right\}$

Now, for the second and third run of the process where we populate the chrominance pixel values, consider Fig. 1.15 (b), $R_2 = \frac{R_1+R_3}{2} + \frac{(-G_1+2G_2-G_3)}{2}$ and $R_4 = \frac{R_1+R7}{2} + \frac{(-G_1+2G_4-G_7)}{2}$. This was the case when the chrominance pixels have neighbors of the same type in the same row or column. In case we have to estimate

the missing R_5 we employ the same method as we did in the interpolation of the green band. In this case,

$$hg = abs(-G_3 + 2G_5 - G_7) + abs(R_3 - R7)$$

and

$$vg = abs(-G_1 + 2G_5 - G_9) + abs(R_1 - R9)$$

then

$$R_5 = \left\{ \begin{array}{ll} \frac{R_3+R_7}{2} + \frac{-G_3+2G_5-G_7}{2} & \text{if hg} < \text{vg} \\ \frac{R_1+R_9}{2} + \frac{-G_1+2G_5-G_9}{2} & \text{if hg} > \text{vg} \\ \frac{R_1+R_3+R_7+R_9}{4} + \frac{-G_1-G_9+4G_5-G_3-G_7}{4} & \text{if hg} = \text{vg} \end{array} \right\}$$

Fig. 1.16 gives the results obtained for the adaptive color plane based demosaicking process. We see that the adaptive demosaicking process gives the best result of all the bands. The RMSE value is the least of all the demosaicking methods, thus making this method the most suitable method for demosaicking color images.



(a) Mosaicked image obtained from Bayer CFA

(b) Adaptive color plan based demosaicked Output

Figure 1.16: Adaptive Color Plane based Demosaicking Results for Color Images

1.5 Contribution of Research

This thesis work concentrates on introducing a new multi-spectral image acquisition system. The Mosaic Focal Plane Array (MFPA) technology which is popularly used in digital color cameras has been used for multi-spectral images.
The major contributions of this thesis work include the extension of existing demosaicking techniques for color images to multi-spectral images. The development of a new demosaicking strategy based on the mosaic filter array formation process. A novel Maximum a-Posterior probability (MAP) based approach that deals with performing demosaicking in the presence of noise and degradations has also been developed.

1.6 Thesis Outline

The outline of the thesis is as follows:

Chapter 1 provided a brief introduction to the work done in the thesis. It also provided a literature survey on the related topics to the thesis work.

Chapter 2 deals with the two main processes of the MFPA technology - the *mosaicking* and the *demosaicking* process. The chapter involves a thorough discussion on the design requirements for the mosaicking process. The discussion leads to the development of a seven-band multi-spectral mosaic filter array using a generic mosaic filter array creation technique. The demosaicking process forms the major part of the chapter. The details of demosaicking process including the design requirements and various interpolation-based demosaicking strategies have been discussed. A development of a new interpolation-based demosaicking strategy has been designed and discussed in this chapter.

Chapter 3 delves into optimizing the demosaicking process. The chapter focuses on treating the demosaicking process as a traditional image restoration problem and solving the optimization problem. A *Maximum a-Posteriori* (MAP) based method has been developed to solve the demosaicking problem. The discussion on the MAP problem involves the possible prior and sensor models that have been used in this thesis. A solution for the MAP problem has been developed using

the Gradient Descent method.

Chapter 4 exhibits the experimental results obtained for the MFPA based methods discussed in the Chapter 2 and Chapter 3. The chapter starts off with a discussion on the experimental process and the use of the multi-spectral image database for experimentation purposes. The use of different metrics for comparing the demosaicking methods have also been discussed. This is followed by the listing of all the results for different methods. The chapter compares all the results and comments upon the best method amongst the numerous methods implemented in this thesis work.

Chapter 5 summarizes the contributions in this thesis work and provides some possible extensions for future work in this area.

Chapter 2

MFPA Technology for Multi-Spectral Images

There are very few human beings who receive the truth, complete and staggering, by instant illumination. Most of them acquire it fragment by fragment, on a small scale, by successive developments, cellularly, like a laborious mosaic.

– Anais Nin

The discussion in Section 1.4 gave an overview of the use of mosaic focal plane array technology in digital cameras. We continue the discussion in this chapter by focussing on the development of demosaicking algorithms for multi-spectral images.

Before going into the details of mosaicking and demosaicking methods for multi-spectral images, we need to understand that the methods used for color images cannot be directly applied to multi-spectral images. New algorithms have to be developed for both mosaicking and demosaicking of multi-spectral images. In case of the mosaic pattern, we use a color filter array (for e.g. Bayer color filter array) in color images. The same array cannot be used in case of multi-spectral images simply because the Bayer color filter array has provisions only for three band images. There is a need for a new mosaic pattern or filter array

that can be used for the multi-spectral images. Another area where there is a need for modification is the demosaicking process. Demosaicking algorithms that have been used for color images are specific only to color images. To develop new demosaicking algorithms or to extend the same algorithms to multi-spectral images we need to first identify the factors that control the demosaicking process. This chapter will give a brief overview of the mosaicking process and the factors that affect the mosaicking process. It will also deal with the extension of the mosaicking process to multi-spectral images. This will be followed by a discussion of demosaicking process for multi-spectral images.

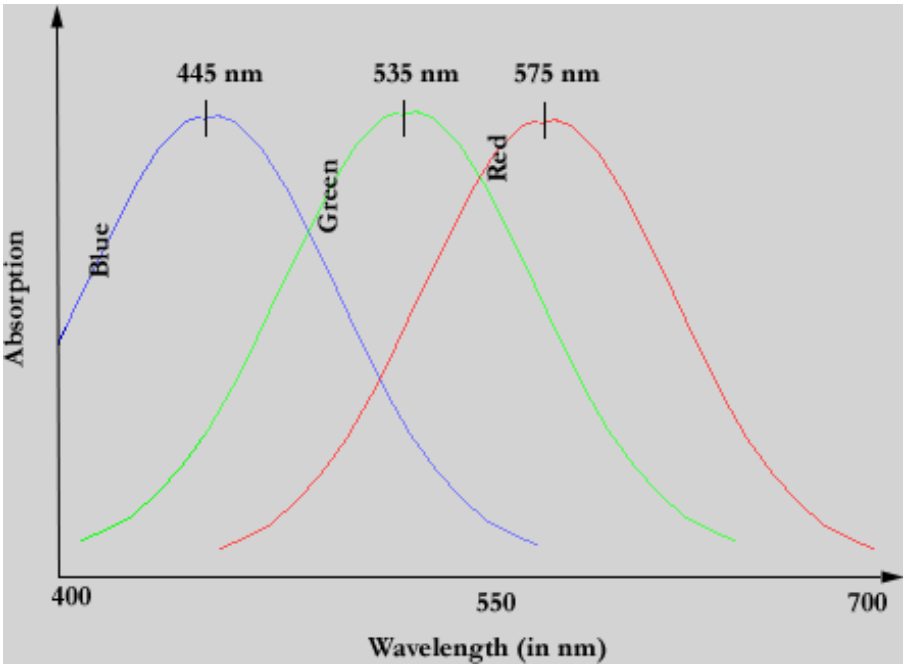
2.1 Mosaicking of Multi-spectral Images

The first thing that comes to mind when we talk about the mosaicking process is the mosaic pattern that is used to mosaic the actual multi-band image. As discussed in the previous chapter, there are many mosaic patterns available for mosaicking color images. The most popular of them is the Bayer color filter array (CFA). However, the use of the Bayer CFA is limited only to color images (i.e, three channel images). There is a need for a mosaic filter array that can accommodate more than three bands per image. Miao *et al.* [Miao et al., 2003] have proposed a mosaic filter array that can be used for multi-spectral images. The paper discusses a generic method to generate mosaic filter array patterns for a given number of spectral bands and their respective frequencies of appearance. Using this method we can generate filter arrays for images with any number of spectral bands. In this section, we will give a brief overview of this method and its applications.

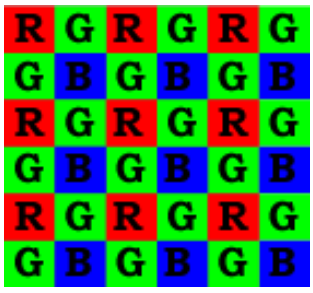
Before we go to the mosaic filter array formation procedure, it is important to review the factors that affect the mosaicking process. These factors play a crucial role in the design of the mosaic filter array. A thorough discussion is given below on the various factors that control the mosaic filter array pattern. The discussion will describe the factors based on color filter arrays and then move on to their extension to multi-spectral mosaic filter arrays.

Frequency or Probability of Appearance The frequency of occurrence of a spectral band plays an important role in the design of the CFA. The factors that affect the frequency of occurrence of a spectral band depend on the purpose for which the camera is being used. For example, in case of digital still cameras, the main purpose is to capture a scene and convert it into a color image. That is, the final output image has three spectral bands (Red, Green and Blue) and the main motive is to form the most visually appealing output. Thus, the human visual system is taken into consideration for the development of a CFA. Extensive research work on the human visual system has concluded that the human visual system is more sensitive to luminance changes than chrominance changes [Mullen, 1985]. This means that for creating a best visually appealing output, i.e. forming the best estimate of the actual scene, the CFA must give more importance to luminance information than the chrominance information. Now, looking at the luminance characteristics of the red, green and blue spectral bands for the human visual system, in Fig. 2.1(a)(*Courtesy [Gonzalez and Woods, 2003]*), we see that the green spectral band is the major contributor of luminance information. Thus, the green spectral band is considered as the luminance band in most of the existing CFAs. The luminance band is sampled at a higher rate compared to the other existing bands. For example, in the Bayer CFA, the green band is sampled with twice the rate compared to the red and the blue spectral bands [Bayer, 1976]. The green band is made to occur at every alternate pixel location as we see in Fig. 2.1(b).

Spectral Consistency In the process of designing a CFA it is important to consider the reconstruction process that follows the mosaicking process. To ensure an uniform reconstruction performance across the whole image, it is important for each pixel to have the same number of neighbors of a certain band within a specific neighborhood. This is an important characteristic of the CFA because the reconstruction process mainly relies on interpolation algorithms. The basis of interpolation algorithms is to gather information about a missing spectral band at a specific pixel location from the neighborhood of the pixel. To ensure consistency in the interpolation process throughout the image, it is important that a specific spectral band shares the same type of neighbors all throughout the image. This

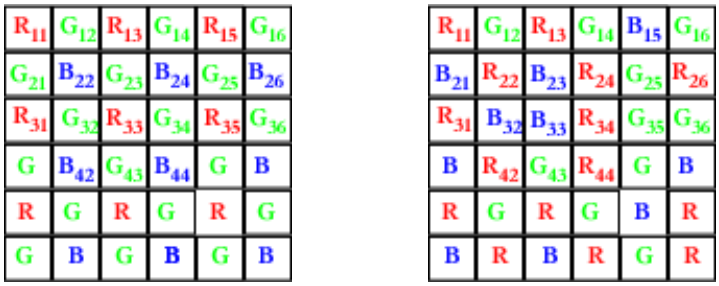


(a) Absorption of light by the human eye for different spectral bands



(b) Bayer CFA

Figure 2.1: Characteristics of the Human Visual System and the Bayer CFA



(a) Bayer CFA (b) Pseudo-random CFA

Figure 2.2: Comparison of Spectral Consistency in CFAs

characteristic of a CFA is illustrated in Fig. 2.2. The Bayer CFA in Fig. 2.2(a) is spectrally consistent. We see that, the pixel with red spectral band information (for example R_{33}) has four green pixels in its four-neighborhood definition (i.e., G_{23}, G_{34}, G_{43} and G_{32}) and four blue pixels in its eight-neighborhood definition (i.e., B_{22}, B_{24}, B_{44} and B_{42}). This is the case with every red pixel all throughout the image. Similarly, the blue pixel has four green pixels in its four-neighborhood and four red pixels in its eight-neighborhood definition. Same is the case with the green pixels. Now, the random CFA in Fig. 2.2(b) does not display spectral consistency. We see that, the type of neighbors shared by each spectral band is not the same all throughout the image. For example, the red pixel R_{22} has three blue (B_{23}, B_{32} and B_{21}) and one green pixel (G_{12}) as its four neighbors, whereas another red pixel, R_{34} has two red and one pixel of each blue and green as its four neighbors. Thus, the random CFA is difficult to interpolate using the traditional interpolation techniques.

Uniform distribution The distribution of pixels throughout the CFA also plays an important role in the design of a CFA. The CFA has to have a uniform distribution of all the spectral bands throughout the image. This makes the reconstruction process more systematic and produces a better result. The Bayer CFA has an uniform distribution of all the spectral bands throughout the image.

Ease of Interpolation As mentioned earlier, the design of the CFA must also take into account the reconstruction process. The design of the CFA should

be such that the reconstruction of the color image should be computationally less intensive and easy to implement. Traditionally, the reconstruction of the color image from a CFA has mainly depended on interpolation algorithms. The ease of interpolation depends on the arrangement of the spectral bands in the CFA. It is difficult to interpolate randomly generated CFAs than to interpolate fixed CFAs. In a fixed CFA (like, the Bayer CFA), since the arrangement of spectral bands is repetitive, the algorithm for interpolation can be easily extended from one pixel to all the pixels of the same type throughout the image. Whereas, in a randomly generated CFA, due to spectral inconsistency, each pixel of the image has to be addressed and interpolated individually. This makes the reconstruction process extremely complex. Methods other than normal interpolation techniques have to be adopted to reconstruct images from a random CFA. Zhu *et al.* [Zhu et al., 1999] have proposed a reconstruction process for random CFAs based on edge detection and boundary interpolation methods. This technique does not require any knowledge about the pattern of the CFA.

Unlike color images, the goal of using the MFPA technology in the case of multi-spectral images is not to obtain a visually appealing output, but to obtain an output which can be used to classify objects in the image. This makes the band selection process in multi-spectral images different than the color images. In a color image, the luminance channel (green band) is given more importance because human visual system is most sensitive to luminance changes that chrominance changes. But in case of multi-spectral images, since the focus is to analyze the characteristics of the object that has been captured in the image acquisition process, we need to choose the spectral band that has the best response coverage. This becomes the band of importance in a multi-spectral image. Applying this logic in the designing process, the spectral band with the best coverage is made to occur most number of times in the MSFA. That is, the probability of appearance of that band is the highest compared to the other bands in the multi-spectral image. The algorithm for the generic method takes care of these design requirements and generates the required MSFA. A checkerboard pattern is used as the starting point for the creation of a mosaic pattern. The choice of a checkerboard pattern (Fig. 2.3) is justified because it inherently possesses the properties of an ideal

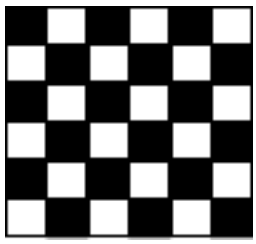


Figure 2.3: Sample Checkerboard Pattern

mosaic filter array. It is symmetric in both horizontal and vertical directions, the black and white blocks are uniformly distributed across the whole image plane and the pattern has the same sampling frequency in horizontal and vertical directions. The algorithm consists of two steps - *decomposition* and *subsampling*. These steps are repeated consecutively until we obtain the required mosaic filter array pattern. In Fig. 2.4 we illustrate the algorithm for obtaining a mosaic filter array pattern for four bands with each band having equal probability of appearance (POA). A binary tree representation is used to illustrate the algorithm. Mathematically, we would like to generate a n-band filter array with each band having a specific probability of appearance (POA). Let the POA for each band be represented by $p_1, p_2, \cdots, p_n$ such that $\sum_{i=1}^{n} p_i = 1$. Now to adjust the POAs of all the bands to add up to 1 we need to ensure that the total number of *leaves* in the binary tree are equal to the total number of bands. Each leaf in a particular level of the binary tree has a $POA = \frac{1}{2^l}$ where l stands for the level in the binary tree. Now, consider the case of the four band filter generation. We see that the binary tree has been divided into two levels and has 4 leaves in total. Each of the band has equal $POA = \frac{1}{2^2} = \frac{1}{4} = 0.25$. That is, the four bands are equally distributed all over the image with a probability of appearance of 0.25 each. The process starts with the checkerboard pattern. The first step is to decompose the checkerboard pattern into two images such that one of them contain only the white pixels and the other contains only the black pixels. For easier representation, we denote the black pixels by 1 and the white by 2. Next, the subsampling step is performed. Each of the level-1 images are subsampled into two more images. The subsampling step, downsamples an image in both the horizontal and vertical directions.

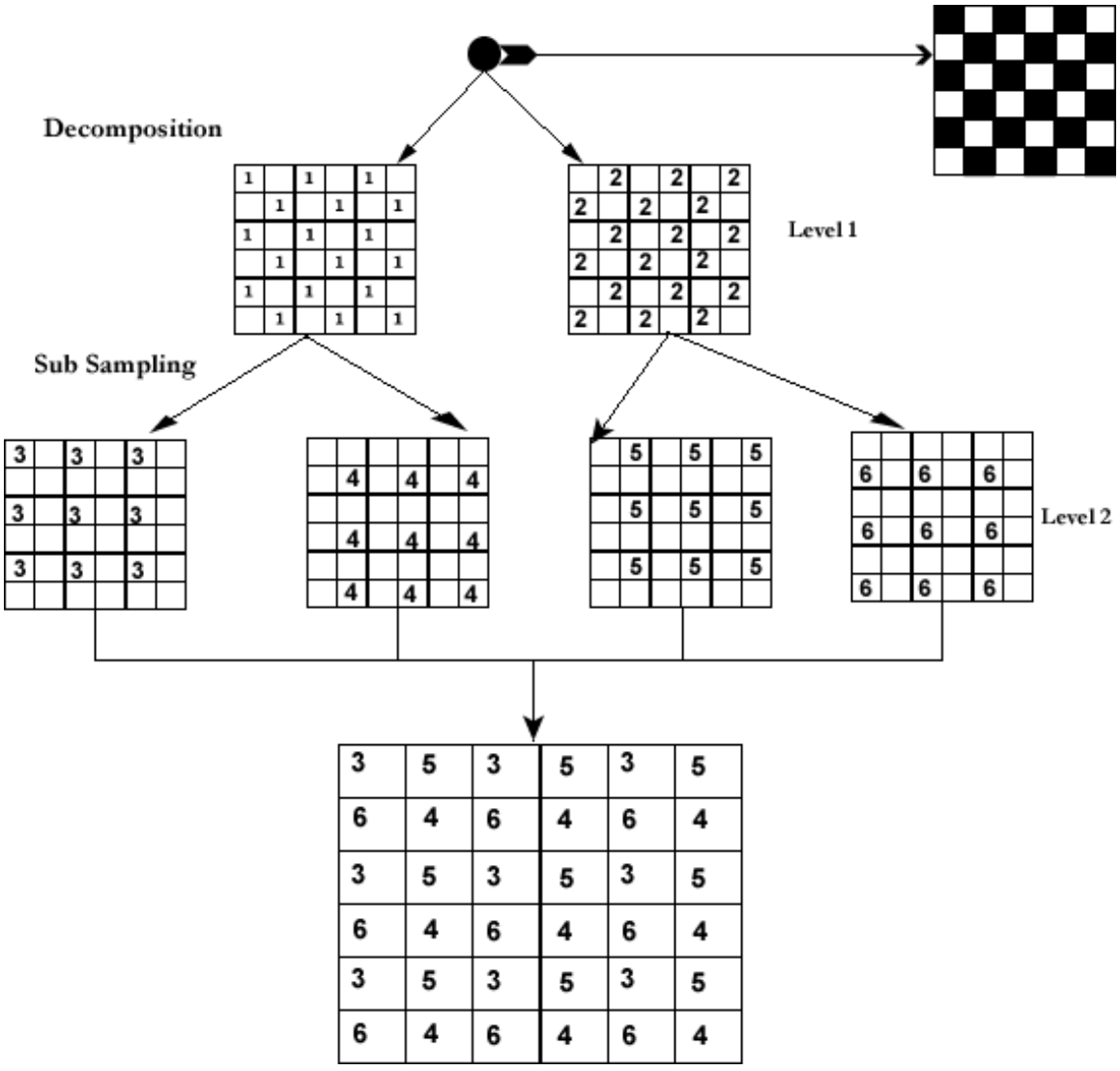


Figure 2.4: Generation of a Four-Band Mosaic Filter Array Pattern

This ensures the pattern to remain spectrally consistent. This step takes us to the next level in the binary tree where four new labels are formed - 3,4,5 and 6. At this step, we check whether we have reached the total number of required bands. In case we have then we combine all the labels in the latest level and form the final mosaic pattern. Otherwise the decomposition and the subsampling steps are repeated until we reach the required number of bands. In the case of a four band pattern, we have already achieved the goal and so we stop at the second level. Now, for the four band pattern to be used as a mosaic pattern it should be able to satisfy all the design requirements for a mosaic filter array. This is verified in Fig. 2.5.

Probability of Appearance Each of the four bands in the filter array have the same probability of appearance (POA = 0.25). This means that all the bands in the filter array are sampled at the same rate. Comparing this with the Bayer CFA, we see that in the Bayer CFA, the green band is doubly sampled than the two chrominance bands. This makes the frequency of occurrence of the green band double than the other two bands. However, in this case the frequency of occurrence for all the four bands will be the same all throughout the filter array.

Spectral Consistency Each spectral band in the pattern has the same type of neighbors all throughout the image. Considering the example given in Fig. 2.5. Two pixels having band 1 information, 1_a and 1_b share the same type of neighbors. This is the case with every band in the filter array. This verifies the spectral consistency of the filter array.

Uniform Distribution By uniform distribution we mean that the occurrence of each of the bands in the filter array is equally distributed all over the filter array. By observation, we can see that the pattern follows an uniform distribution.

The four band pattern we have generated using the generic method resembles the mosaic pattern used by Sony in its digital cameras. It is called the Sony *RGBE*

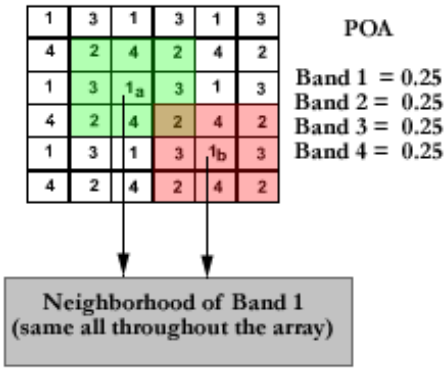


Figure 2.5: Four-Band MSFA

mosaic pattern, which uses the usual three bands - red, green and blue, and an additional band called the emerald band to mosaic color images.

In context of the work done in this thesis we will be using multi-spectral images which have seven bands. We used the generic method to develop a MSFA for seven bands. Fig. 2.6 shows the steps followed in the creation of the seven band MSFA using the generic method. We see from the figure that the algorithm requires one of the seven bands to have twice the POA than the rest of the bands. Band-3 in this case, occurs twice as frequently than the rest of the bands in the MSFA. The band has been sampled with twice the sampling rate than the rest of the bands. This assigns more prominence to that band than the rest of the bands. Usually, the band with the highest coverage is made to have the highest POA. To analyze the seven band MSFA even further, consider Fig. 2.7. From the figure, a neighborhood of a band 1 pixel is considered. In this case, we consider a 5×5 neighborhood. The size of the neighborhood is important when demosaicking is being performed on the MSFA. We can see that the MSFA satisfies all the design requirements for a mosaic filter array. It has uniform distribution of all the 7 spectral bands, it is spectrally consistent (for eg., observe the neighborhood around all the band-1 pixels. The neighborhood is consistent all throughout the array) and the probability of appearance of each band is also justified according to the generic algorithm (band 1 has twice the sampling rate compared to the rest of the bands and thus it appears twice more number of times than the rest of the

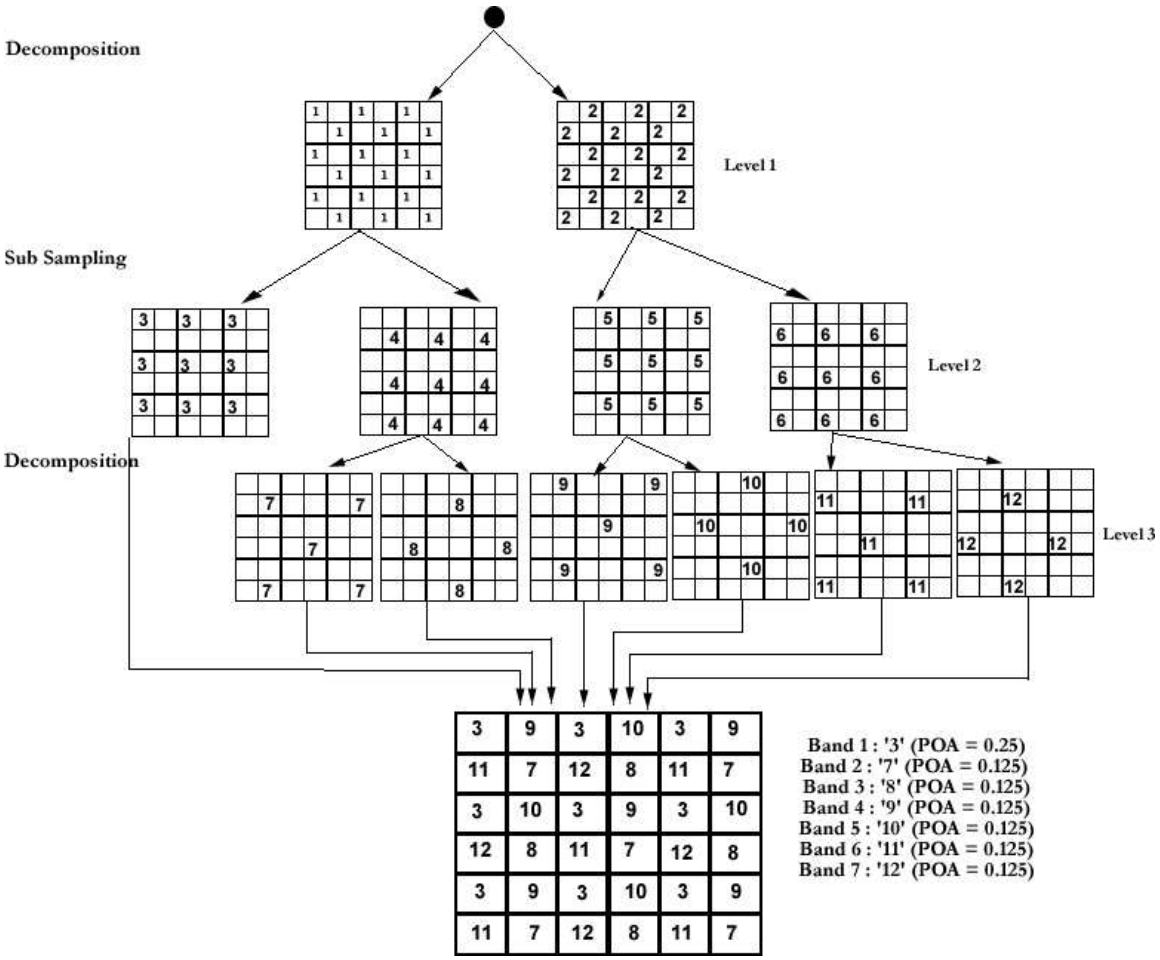


Figure 2.6: Seven-Band MSFA Generation

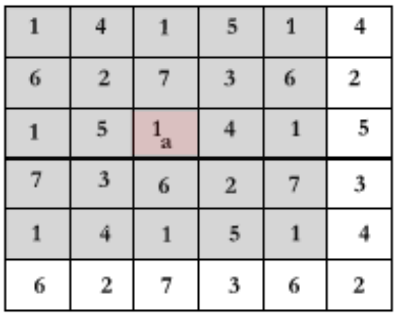


Figure 2.7: Sample Seven-Band MSFA

bands).

2.2 Demosaicking of Multi-Spectral Images

In the previous section, we analyzed ways of creating filter arrays for multi-spectral images. The next step in the MFPA process is to reconstruct the multi-spectral image from the mosaicked image. Before we venture into different aspects of demosaicking multi-spectral images, we have to understand the major difference between demosaicking color and multi-spectral images. In Sec. 1.4.2 we analyzed some popular demosaicking techniques for color images. These methods were restricted only to color images as they use the special characteristics of color images like prominent luminance band, correlation of the luminance and chrominance bands, etc. These features may not exist in multi-spectral images. So there is a need to make changes to the algorithms such that we can apply them to multi-spectral images. To make the required changes in the algorithms we need to understand the factors that affect the demosaicking process. This understanding will enable us to develop modification in a proper direction.

2.2.1 Factors Affecting the Demosaicking Process

The major factors that affect the demosaicking process are listed below.

Mosaic Filter Array Pattern It is obvious that the design of the mosaic filter array pattern that is being used to sample the multi-band image affects the reconstruction process. It is necessary to have complete knowledge about the mosaic pattern before developing the demosaicking algorithms because the mosaic pattern reveals important information about the number of bands present in the actual image and the relationships that the bands share between each other. This is illustrated in Fig. 2.8. The figure displays three mosaic filter arrays with different number of spectral bands. Each of the MSFA reveals important information about the total number of bands present in the actual image, the neighborhood information of each of the band, the sampling rate of the spectral bands and also the distribution pattern of spectral bands all throughout the image. The knowledge of neighborhood information for each pixel is necessary because the demosaicking process uses neighborhood information to interpolate missing band values at each pixel location. Similarly, the knowledge of the rate at which each band is sampled affects the order in which the missing bands are estimated. Due to the existence of greater number of samples, the highest sampled band is usually estimated before any other spectral band. The manner in which the spectral bands are distributed in the mosaicked image also plays an important role in designing the demosaicking algorithms. Most of the MSFAs are uniformly distributed and are spectrally consistent; this makes recurring patterns appear throughout the MSFA. If these recurring patterns are known, the demosaicking algorithms can be developed to consume less computation time and also be more effective. For example, consider the shaded portion in the 3-band MSFA in Fig. 2.8. We see that the pattern is repeated all throughout the image array. The pattern is such that taken any four adjacent pixels, one of the diagonal will have band 1 pixels and the other diagonal will have one pixel each of the rest of the two bands. This information is important because, if the demosaicking algorithm is embedded with this information then the reconstruction process can be made faster and probably more effective.

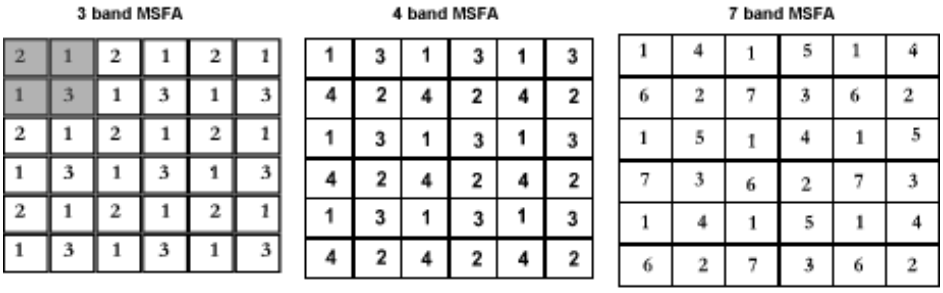


Figure 2.8: MSFAs for Different Number of Spectral Bands

Number of Spectral Bands Multi-spectral images are characterized by the total number of spectral bands, type of spectral bands and their size. The number of spectral bands in the image affects the design of the mosaic filter array which in turn affects the demosaicking process. The main effect of the increase in the number of bands, on the MSFA, is that it makes the distribution of missing bands in the MSFA even more complicated. The spread of missing band values increases as the number of spectral bands increase. This makes the interpolation process more complicated and less accurate. In Figure 2.8, we see that the distribution of missing pixels is less complex in the 3-band MSFA than the 7-band MSFA. The demosaicking algorithm has to consider a bigger neighborhood in the case of the 7-band MSFA so that the neighborhood will have a proper representation of all the spectral bands.

Neighborhood Considerations Neighborhood of a pixel plays a very important role in the design of demosaicking procedures [Ramanath et al., 2002]. It may be expected that by increasing the size of the neighborhood, the estimate of missing pixel values can be done more accurately. But this is not true because of two reasons. Firstly, as the neighborhood size increases the relationship between the missing pixel values and the existing pixel values in the neighborhood deteriorates. Thus, the probability of estimating the exact missing pixel value from its neighborhood pixels decreases to a great extent. Secondly, an increase in the

neighborhood size makes the demosaicking method computationally expensive. However, on the other hand, a smaller neighborhood does not guarantee better results. Thus, a trade-off must be met with respect to neighborhood considerations such that appropriate missing pixel value can be estimated and also the method is not too computationally expensive.

Edge Information Demosaicking is performed by interpolating the missing pixel values based on the neighborhood pixel values. An average of the neighboring pixel values may be a good estimate for the missing pixel value only in areas of the image which contain uniform pixel values. By uniformity, we mean the intensity values do not change drastically from pixel to pixel. But in the case of an occurrence of an edge, the pixel values undergo a sudden change in which case the average of the neighboring pixel values of the edge pixel will not be a good estimate. In such cases, the edge direction and the strength of the edge should be considered while interpolating the missing pixel values. The direction of interpolation plays an important role in interpolating missing pixels. In case of an edge, the direction of interpolation has to be chosen such that it matches the gradient direction. The gradient direction is perpendicular to the edge direction. This will ensure that an originally occurring edge will be retained and thus the estimate will be more appropriate. We have seen these kinds of edge information considerations in some of the advanced demosaicking methods for color images in Sec. 1.4.2.

Inter-Band Dependence Some of the existing demosaicking methods for visual band images have been derived from the fact that the edge information in one band is related to the edge information in another band. In such cases, the demosaicking process may be allowed to use one band information to interpolate missing values of another band. This type of estimate would be totally valid since both the bands have common edge information. For example, in visual band images it has been found that the R, G, B bands have almost common edge information. Thus, G being the highest sampled band may be interpolated first and then the interpolated values may be used to interpolate the missing values of B



Figure 2.9: Example of Inter-Band Dependence in Images

and R bands. This has been illustrated in Fig. 2.9. We see that all the three bands almost have the same edge information. But in multi-spectral images, since there are spectral bands other than the visual spectral bands, the inter-band relationship may not be the same as in color images. However, if it can be determined as to whether there is any existing correlation between any two bands then the band information may be used for other band's interpolation process.

Practical Issues Practical issues like the time taken for the demosaicking process also have to be considered. The demosaicking process is the most time consuming process when compared to data capturing and mosaicking processes. The time taken for the demosaicking process should be minimized by use of proper programming techniques so that it would be practical to implement it in a real-time

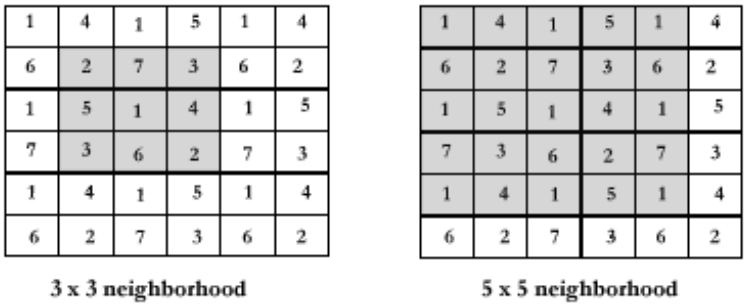


Figure 2.10: Different Sized Neighborhoods in the Seven Band MSFA

environment.

2.2.2 Demosaicking Methods for Multi-Spectral Images

Now that we have seen the design aspects of the demosaicking process, it is time we go through some of the demosaicking techniques that would be used for multi-spectral images. Most of the demosaicking methods that are discussed in this section, are largely based on the existing demosaicking methods for color images. Required changes have been made so that they comply with a different MSFA. Each method will be discussed in detail with a prediction of how the result will look after performing demosaicking. The results for these methods are given in Chapter 4.

Bilinear Demosaicking We start the discussion on demosaicking methods with the most simple and basic method called the Bilinear demosaicking method. Bilinear demosaicking focuses on interpolating the missing pixel values from its neighboring pixel values. The missing pixel value is found by averaging the existing pixel values of the same spectral band in the neighborhood of the pixel. The key to this method is the choice of the neighborhood size. From Fig. 2.10 we see that a 3×3 neighborhood does give a proper representation of all the spectral bands in the image. Moreover, the distribution of the spectral bands in the neighborhood is also not consistent. On the other hand, a 5×5 neighborhood gives a clear representation of all the spectral bands in the image. However, the

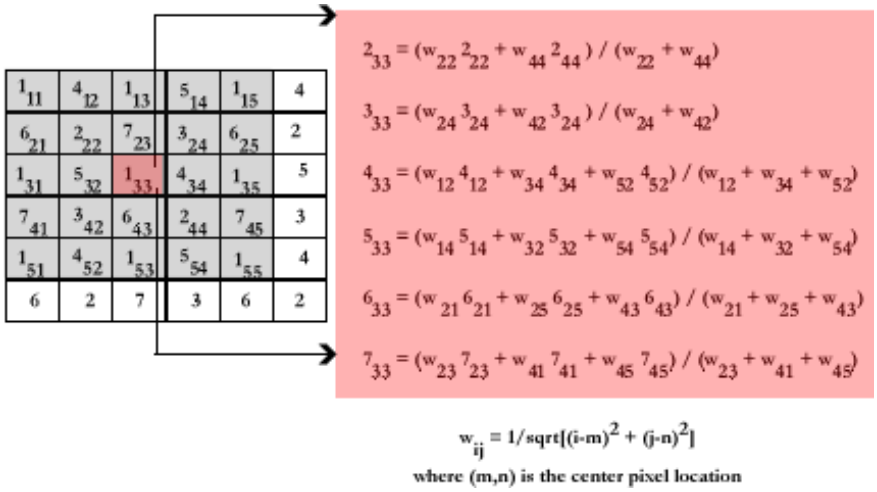


Figure 2.11: Bilinear Demosaicking of a Seven Band Multi-spectral Image

problem with increasing the neighborhood is that the pixels are displaced by a greater distance from the center pixel. That is, the effect of pixel 7_a will be more on the center pixel than the other two 7^{th} band pixels - 7_b and 7_c. This discrepancy can be taken care of by assigning weights to each of the pixels and then summing them up. The weights assigned to each of the pixel is inversely proportional to the distance between the current and the center pixel. A detailed illustration of the bilinear demosaicking process can be obtained from Fig. 2.11. In general, the estimate of a missing pixel value is given by Eq. 2.1

$$B_{mn} = \frac{\sum_k \sum_l w_{kl} B_{kl}}{\sum_k \sum_l w_{kl}} \tag{2.1}$$

where $w_{kl} = \frac{1}{\sqrt{(k-m)^2+(l-n)^2}}$, B is the band to be interpolated and (m, n) is the pixel that is being interpolated. In the discussion of bilinear demosaicking for color images (Sec. 1.4.2) we were considering a 3×3 neighborhood and there was no need for adding weights to each of the pixels. This was because the 3×3 neighborhood gave an appropriate representation of all the spectral bands and also had equidistant neighbors around the center pixel. We also observed that the bilinear demosaicking method produced artifacts at the edge locations called the *zipper effect*. This effect also occurs in the reconstruction process of the

multi-spectral images. This can be seen in chapter 4 where all the results for the demosaicking methods for multi-spectral images are listed.

Cok's Demosaicking Method This demosaicking method is based on the constant-hue based demosaicking method invented by Cok for color images. As we saw in the previous chapter, the method focuses on the removal of edge artifacts formed by the bilinear demosaicking method. Cok [Cok, 1987] suggests that the occurrence of the artifacts in the process of interpolation is due to sudden changes in hue in that image region. Hue was defined by the ratio of the chrominance and luminance band values ($\frac{R}{G}$ and $\frac{B}{G}$) at a particular pixel location. Revisiting the algorithm listed in the previous chapter, we find that in case of non-uniform hue images, the hue is maintained by taking a small neighborhood into consideration and interpolating the hue values based on the neighborhood information.

We try to apply this method to the multi-spectral images. Before we start extending the method to multi-spectral images, we need to realize that unlike color images, the concept of hue does not exist in the case of multi-spectral images. This is simply because multi-spectral images have spectral bands which may not be part of the visual spectrum (for instance, R, G and B may not be the only spectral bands in the multi-spectral image). Still taking the main essence of the method, we could extend the logic by considering the main band (i.e., the band with the highest coverage) as the luminance band and the rest of the bands as the chrominance bands. This analogy will only work if there is a relationship between the main band and the rest of the bands in the multi-spectral image like the one shared by the luminance and the chrominance bands in color images. One way to check if this method works is to actually mathematically calculate the inter-band correlation between each of the bands in the multi-spectral image and then analyze to the correlation values to come up with a valid conclusion on whether the extension of the method to multi-spectral images is justified or not. Fig. 2.12 illustrates Cok's demosaicking method for seven band multi-spectral images. The method is a two pass process. In the first pass, the main band is interpolated. The interpolation is performed by using the usual bilinear interpolation method.

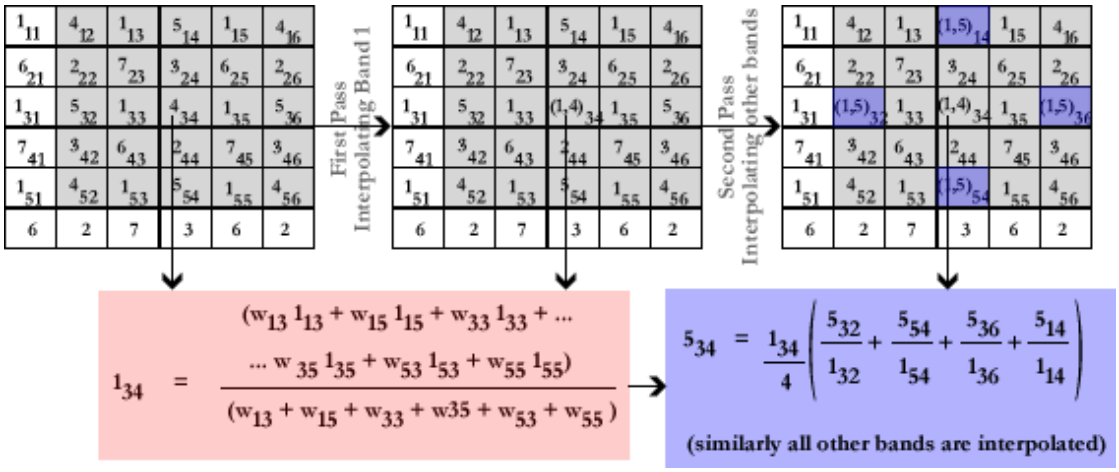


Figure 2.12: Cok's Demosaicking of a Seven Band Multi-spectral Image

In the next step, all the other bands are interpolated as given in Eqn. 2.2.

$$B_{ij} = M_{ij} \sum_{(a,b)\varepsilon\aleph} \frac{B_{ab}}{M_{ab}} \tag{2.2}$$

where B is the band to be interpolated at pixel location (i, j) and M is the main band (in the case of our seven band images, M is band 1). $\aleph$ signifies the neighborhood of the current pixel (i, j) and the pixels $(a, b)\varepsilon\aleph$ (for eg., the grey shaded part in the Fig. 2.12 represents the 5×5 neighborhood ($\aleph$) of pixel location $(3, 4)$).

Median Based demosaicking method This method is inspired from the median based demosaicking method patented to Freeman. The original method [Freeman, 1988] was developed for color images mosaicked using the Bayer CFA. We will try to extend the method to the seven band multi-spectral images. The method is focuses on improving the edges in the bilinear demosaicked image. The green band, i.e. the luminance band, in a color image is interpolated first and then the rest of the two chrominance bands are interpolated with the help of a median filter. The interpolation of the chrominance bands is performed by taking two difference images, $R-G$ and $B-G$, and then applying a median filter to the images. Then the

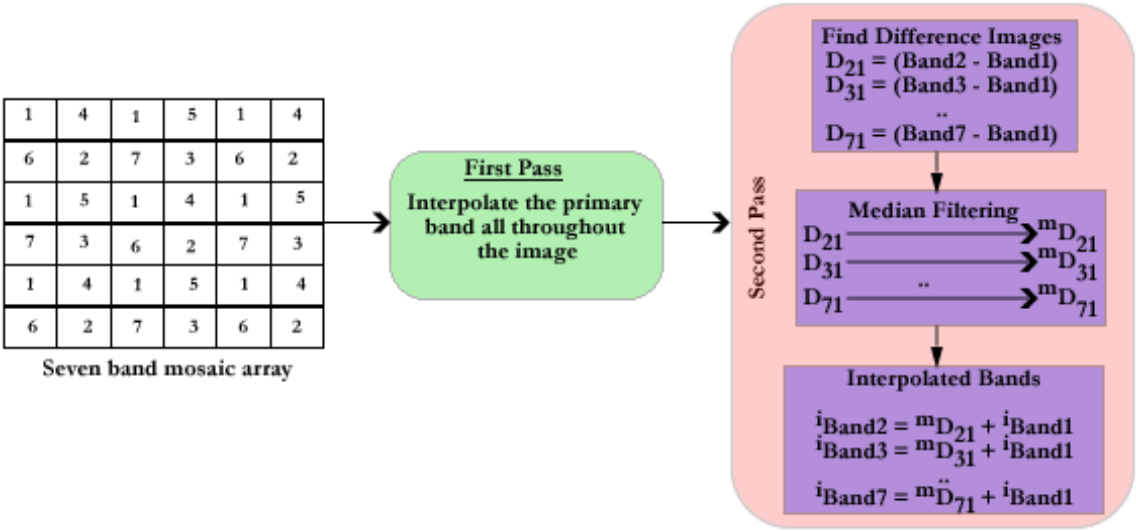


Figure 2.13: Median Based Demosaicking for Seven Band Multi-spectral Images

interpolated luminance image is added to the median filtered difference images to form the whole color image.The method has been explained in detail in Sec. 1.4.2. In case of multi-spectral images, we use the primary band (the band with the highest coverage) in place of the luminance band. The first pass will interpolate the primary spectral band values throughout the image. The second pass will perform median filtering on the difference images and then add the result to the interpolated primary band image. The two steps involved in median filtering are illustrated in Fig. 2.13. The median based demosaicking method performs well only in images where all the spectral bands share the same edge information. This seems to be true in the case of color images. However, this may not be true in the case of all the multi-spectral images, thus the performance of this method for multi-spectral images is not guaranteed.

Modified Bilinear Demosaicking The demosaicking methods discussed earlier in this section, are extensions of the methods applied for color images. Each of the methods has inherent problems which can affect the output of the demosaicking process. The bilinear demosaicking method, for instance, does not consider edge information and thus falls prey to edge-related artifacts called the *zipper ef-*

fect. On the other hand, the Cok's demosaicking method focuses on balancing the changes in hue along the image plane, which happens to be a specific property of color images. Though we could extend the method to multi-spectral images, there is a very small chance for it to produce a perfect demosaicking output. Similarly, the median based demosaicking method, uses edge information specific to color images for interpolating missing bands. There is therefore a need for a demosaicking method that will overcome the drawbacks of the existing demosaicking methods.

The modified bilinear demosaicking method, as the name suggests, is a modification of the usual bilinear demosaicking method. This method has been derived as the reverse process of the mosiac filter array formation process. We start with the MSFA at the bottom of the binary tree (see Fig. 2.6) and at each level try to interpolate the bands that are children derived from another band from the previous level. For example, we can see from the figure, at the last level (i.e., level 3) we have bands 2 through 7 as children from the previous level (i.e. level 2). We start interpolating, bands 2 and 3; bands 4 and 5 ; and bands 6 and 7 at each others positions. Then we move to the one level upwards and treat the band pairs formed in the previous step as one band and then interpolate in the same manner until we reach the start of the binary tree. By the time we reach the top most level of the binary tree we will have an estimate of all the spectral bands at all the pixel positions in the image plane. This is the basis of the modified bilinear demosaicking method. For ease of implementation, this method can also be looked at from another perspective. The method focuses on maximizing the probability of estimating the appropriate missing pixel value by taking variable neighborhood sizes into consideration. As discussed earlier, the neighborhood of a pixel plays a very important role in the interpolation process of missing pixels. The size of the neighborhood should be small so that the chance of finding the exact missing pixel value by interpolation is increased. On the other hand, the neighborhood should be big enough so as to have an adequate representation from all the spectral bands. The method is divided into four passes. These four steps are shown in Fig. 2.14. The first pass focuses on interpolating the primary band all over the image plane using a 3×3 neighborhood (a 3×3 neighborhood is the

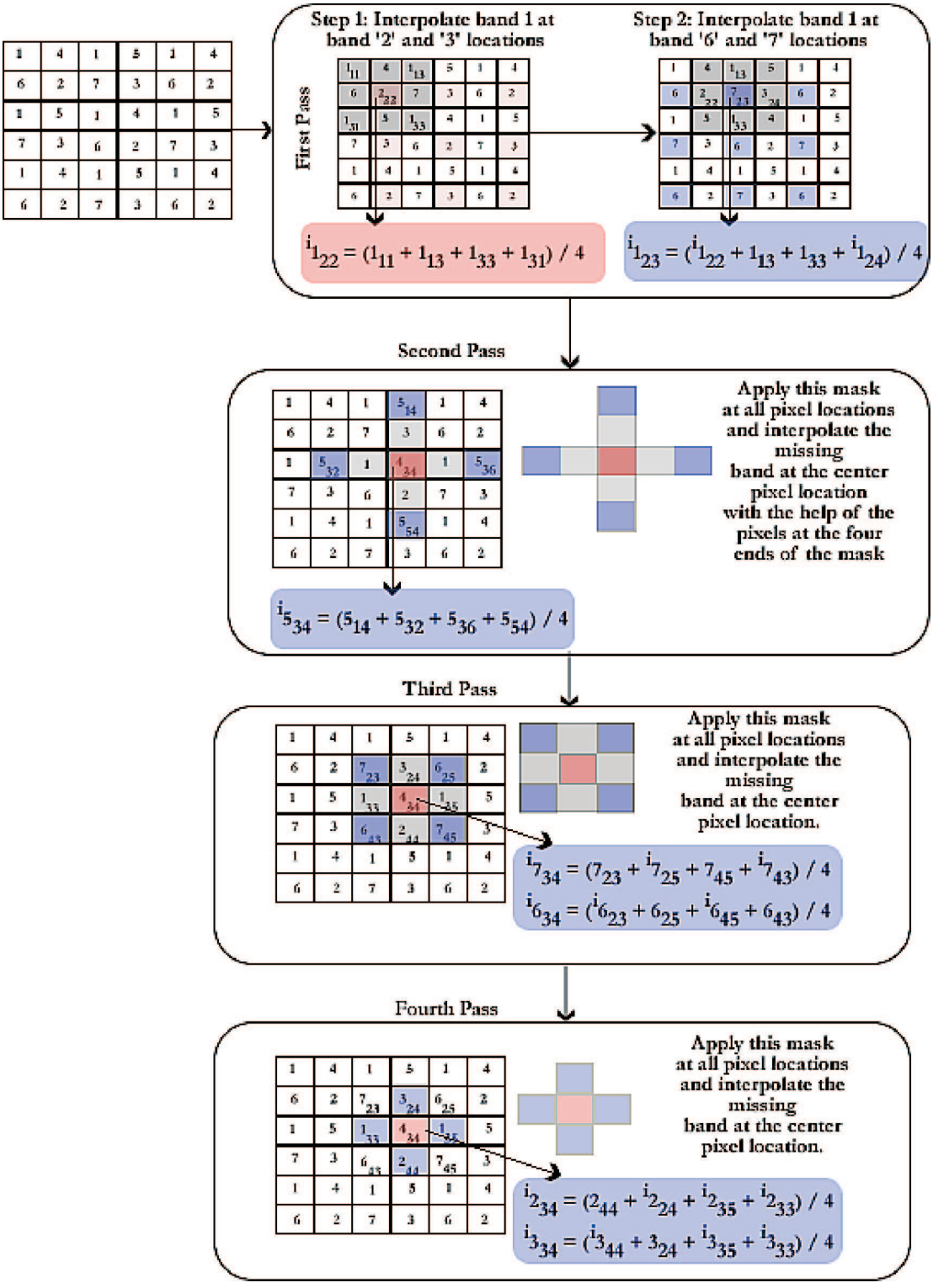


Figure 2.14: Modified Bilinear Demosaicking Process

smallest possible neighborhood with odd number of pixels). The first step, interpolates at locations that have four primary band pixels in their 4-neighborhood definition space. This is followed by the interpolation at locations which have four primary band pixels in their 8-neighborhood space. By the end of the first pass of the method, the primary band would be populated at all the pixel locations in the image. In the following steps, we move on to the interpolation of the rest of the bands in the image. The interpolation is done in a systematic fashion by choosing specific band pairs and then interpolating a missing band at the other band location. This process of choosing band pairs and interpolating them at each others locations acts as a reverse of the decomposition process in the mosaic filter array formation process. In the second pass of the method, band 2 and 3; bands 4 and 5 and bands 6 and 7 are interpolated at each other's location. Due to the uniformity and systematic arrangement of the bands in the mosaic filter array, we are able to form a specific shaped neighborhood for the interpolation of a band at its pair's location. This mask is moved throughout the image plane and the interpolation is performed. By the end of this pass, we have all the pixel locations populated with the primary band and we have all the bands populated at their corresponding pairs' locations. In the next pass, we move to the next upper level of the binary tree in the MSFA formation process. The three band pairs are now treated as three independent band labels, and are paired up with the other labels on the same level depending on how they are formed in the binary tree. The same process of interpolation at the band pair's locations is followed until we reach the top most level of the binary tree. By the time, we have traversed to the topmost level of the binary tree we would have interpolated all the spectral bands at all the pixel locations in the image.

The beauty of this method lies in the fact that it is modeled directly on the mosaic filter array formation process. Additionally, the use of smallest neighborhoods at each step ensures that the estimation process is more appropriate. Theoretically, due to all its design considerations, this method should produce results that are better estimates of the actual scene than the usual bilinear demosaicking method. We will evaluate the actual difference between the two methods in the results chapter (chapter 4) of this thesis.

Modified median based demosaicking method This method is a simple extension of the modified bilinear demosaicking method. The modified bilinear demosaicking method produces an output which has a non-uniformity in the intensity values of the image. This non-uniformity in pixel values generates discrepancies in the image representation. This can cause distortion of the shape of the objects in the image making a system difficult to identify the object of interest in the image. There is therefore a need to improve the uniformity of pixel values throughout the image plane. This is brought about by blurring the image using a median filter. The median filter operates on a small neighborhood and assigns the median value of the neighborhood pixels to the center pixel. This kind of operation throughout the image plane brings about a uniformity in the intensity values in the image.

2.3 Problems With Interpolation Based Demosaicking Strategies

Most of the interpolation based methods discussed in this chapter have been based on traditional demosaicking methods used in color digital cameras. Some of the demosaicking methods have also been used for commercial purposes and have stood the test of time with regards to their performance. However, the design process of these interpolation based methods is heuristic in nature. The development of each of the method is based on improving the existing method by adding a logical extension that will take care of drawbacks of the existing method. There is no concrete mathematical model which can be used to develop an optimal demosaicking process. This proves to be a hurdle in the flexibility offered by the demosaicking method. There is also need for consideration of practical factors like degradation and noise that can be added in the process of image acquisition. The existing interpolation based methods would fail to produce a satisfactory result in the presence of noise and blur added by the acquisition process. There is a need for a method that can take care of these practical aspects of the demosaicking process. These concerns are addressed in the next chapter of this thesis. Chapter 3

would discuss ways of performing the demosaicking process in the presence of degradations and noise. The development of the demosaicking process as an image restoration problem is also considered in the chapter.

Chapter 3

Demosaicking Multi-Spectral Images Using MAP

. . . so-called art restoration is at least as tricky as brain surgery. Most pictures expire under scalpel and sponge.

– Alexander Eliot

The discussion in the previous chapter focussed on the design of demosaicking methods for multi-spectral images. Some of the methods have been inspired from popular demosaicking methods for color images. The main essence of these methods is the use of different interpolation strategies which take various factors like the size and type of neighborhoods, the restoration of edge information, etc to reconstruct the multi-spectral image. The methods however are heuristic in nature. They are based on many assumptions about the characteristics of the image. These assumptions have been made based on logic and extensive experimentation. Though, the methods tend to work and have gained themselves the credit of being applied in commercial applications, still they do not quite allow a great amount of flexibility in their design. One area where the interpolation-based methods fail to function is when there is an occurrence of noise in the image acquisition process. The existence of degradation severely affects the mosaicking process which in turn creates a degraded mosaicked image. This noisy and blurry

mosaicked image is used for forming the demosaicked output. The output then is bound to be severely degraded. Another concern about the interpolation-based demosaicking methods is that they tend to introduce some amount of blur themselves to the reconstructed output. The occurrence of blur is due to the averaging operation that is performed over the neighborhood for populating missing pixel values. There is therefore a need for a method that can take care of the various degradation that creeps into the demosaicking process. This chapter will look into developing such a method that uses image restoration strategies to perform demosaicking that can be used in presence of degradations during the image acquisition process. The next section will focus on explaining how a demosaicking problem can be viewed as a traditional image restoration problem. Next, some of the popular image restoration strategies are discussed. The final section will explain the MAP method for demosaicking multi-spectral images.

3.1 Demosaicking as an Image Restoration Problem

Demosaicking is the reverse process of the mosaicking process. The demosaicking process tries to reconstruct a multi-band image from some samples registered on the mosaic focal plane array during the mosaicking process from the original multi-band image. From Fig. 3.1 we see that the mosaicking process registers only one spectral band at each of the pixel locations in the image plane and the demosaicking process focuses on populating each pixel in the image with all the spectral bands. Looking closer into the demosaicking process, we see that the function of demosaicking is to take a gray scale mosaicked image into consideration and reconstruct a multi-band image out of it. That is, the key function is to reconstruct an image from an imperfect image. This brings us closer to the thought that the process may be considered as a traditional image restoration problem. The traditional image restoration problem restores an image from its degraded version. Fig. 3.2 describes a standard image restoration problem. The original image undergoes degradation due to blur by the averaging process in the

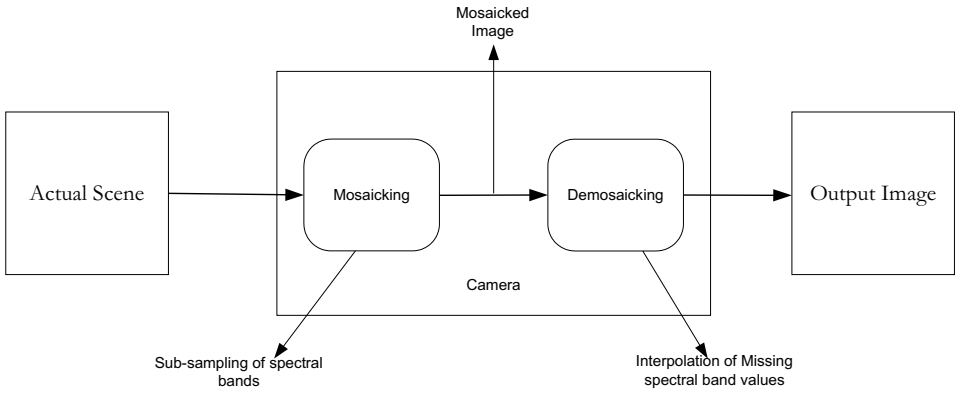


Figure 3.1: MFPA Block Diagram

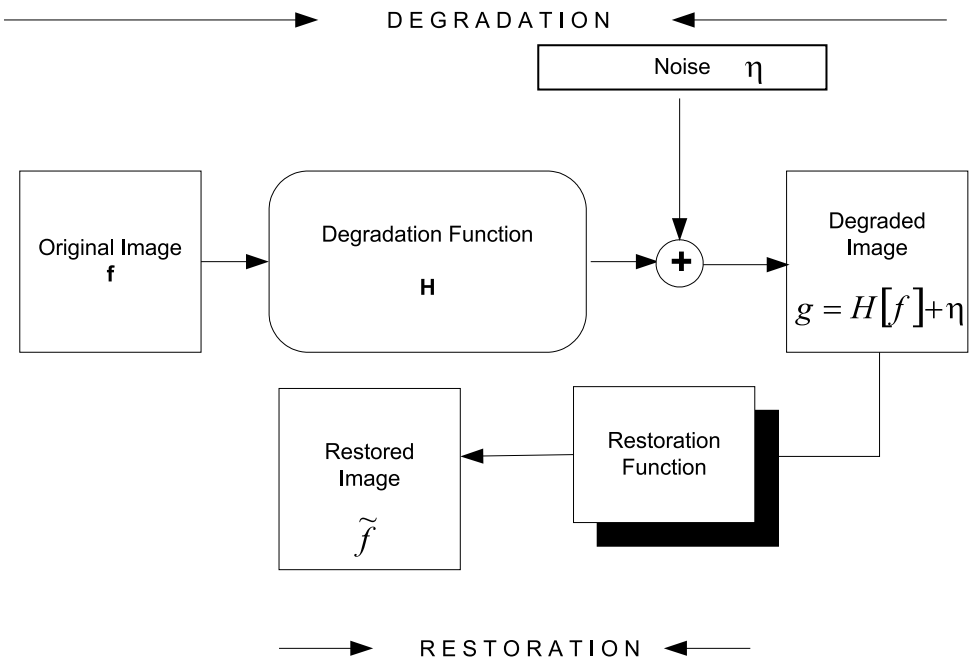


Figure 3.2: General Block Diagram of an Image Restoration Problem

degradation technique and additive noise. The degraded image is then given as an input to the restoration block. The restoration process tries to restore the image by using some restoration filters. The restoration filters are designed based on prior knowledge about the kind of degradation and noise in the degraded image. The problem becomes even more complicated, if there is no prior knowledge about the degradation. Predictably, in real-time situations the prior knowledge about the degradation is seldom known. Now, to model the demosaicking process in the same lines as the image restoration model, we need to identify and relate the concepts of degradation and degraded image. Once we are successful in identifying these terms in the demosaicking process then it is a matter of using one of the standard image restoration procedures to perform demosaicking. Another way of looking at the demosaicking process as an image restoration problem is the case where there is an actual occurrence of blur and noise due to the equipment. In such a case, the demosaicking process is bound to produce a degraded output. This is made clearer from Fig. 3.3. The original image undergoes degradation due to various reasons and then gets noisy due to the presence of additive noise. This degrades the image and forms a degraded image. This degraded image is then fed as input to the mosaicking process. Since the mosaicking process is just a simple subsampling operation over a multi-spectral image, the degradation effect still remains. This will result in the formation of a degraded mosaicked image. This is then given to the demosaicking process, which interpolates the missing pixel values to reconstruct a multi-spectral image. All throughout the process, the degradation that affected the original persists and thus forms a noisy and degraded output. The traditional demosaicking methods have no provision for the removal of the degradation and noise. The only option left in the case of the interpolation-based demosaicking methods is to allow the whole process to continue as usual and then applying multi-channel image restoration methods to clean off the degradation and noise from the demosaicked output to form a cleaner image. This procedure looks cumbersome and may take a long processing time. Instead, if the restoration block is built such that the restoration happens in the process of the demosaicking operation, then the method may be less cumbersome and may also offer greater reliability and control over the whole image formation process. However, having

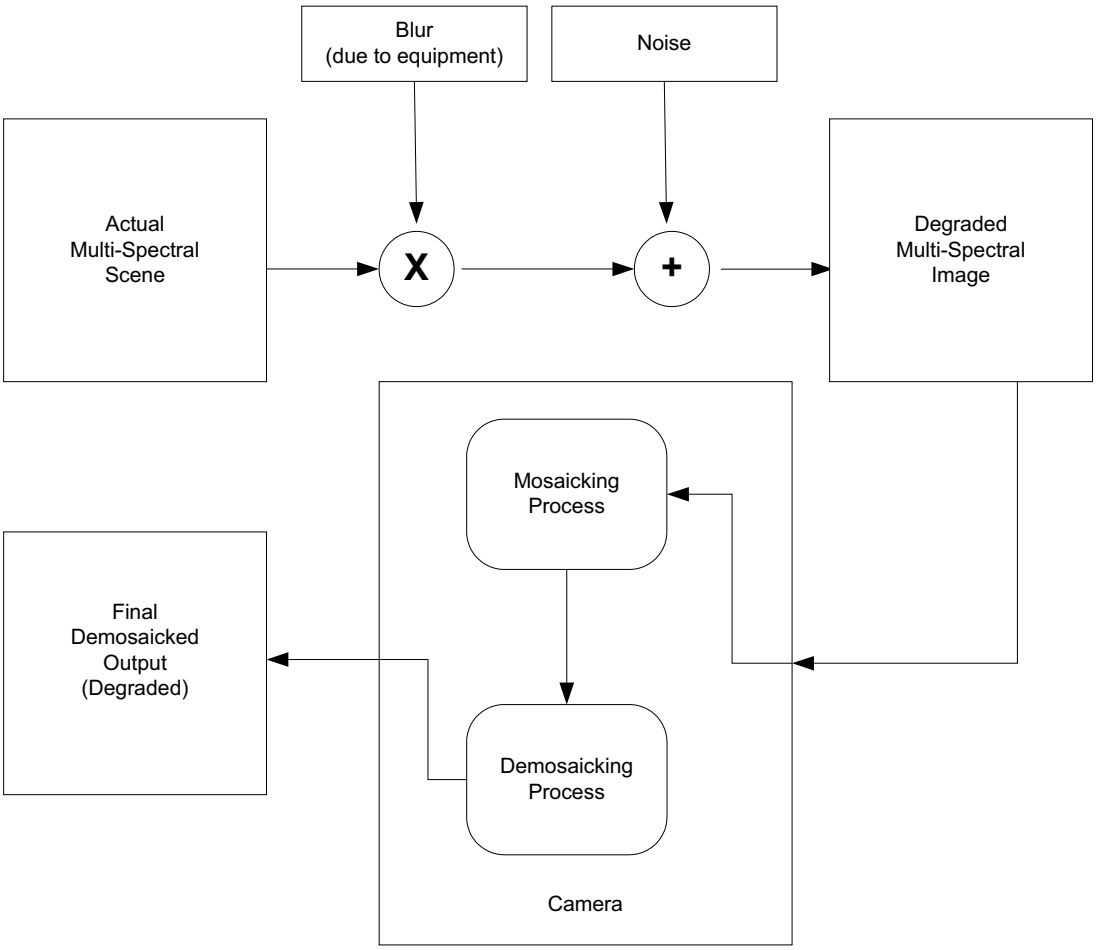


Figure 3.3: Demosaicking In Presence of Degradation

said that the problem of viewing demosaicking as an image restoration problem still remains. The translation will require in depth knowledge about the standard image restoration operations. The understanding will help us relate each block in the restoration process to the demosaicking process. The following section reviews some important concepts in image restoration.

3.1.1 Theory on Image Restoration

Image restoration is one of the most popular research areas in the field of image processing. The objective of the image restoration procedures is to estimate the original image given some prior information about the degradation function and the noise that have corrupted the original image. Looking at Fig. 3.2 we realize that the restoration operation is done by creating restoration filters that will reverse the corruption process done by the degradation function and the additive noise [Gonzalez and Woods, 2003]. The model given in the figure can be modelled mathematically as in Eq. (3.1). This is the spatial representation of the degradation process. The original image $f(x, y)$ is degraded by a linear spatial-invariant blur $h(x, y)$ with some addition of noise from the additive noise term, $\eta(x, y)$ to form the degraded image $g(x, y)$. Now the function of the restoration process is, given the degraded image $g(x, y)$ and some prior knowledge about the blur $h(x, y)$ and noise $\eta(x, y)$ form an estimate of the original image, $\tilde{f}(x, y)$.

$$g(x, y) = h(x, y) * f(x, y) + \eta(x, y) \tag{3.1}$$

The frequency domain representation of Eq. (3.1) is given in the Eq. (3.2). G, H, F and N are the frequency domain counterparts of the observed image, linear spatial-invariant blur, the original image and the additive noise term respectively. The advantage of using the frequency domain representation is that it converts the *convolution* term as a multiplication term, which is easier to operate upon.

$$G(u, v) = H(u, v) \cdot F(u, v) + N(u, v) \tag{3.2}$$

Now, the restoration process is to reverse the degradation process to as high accuracy as possible. Mathematically, the restoration process may be given as in Eq. 3.3.

$$\tilde{f}(x, y) = R[g(x, y)] \tag{3.3}$$

where R denotes the restoration filter and $\tilde{f}$ stands for the final restored image which is an estimate of the original image. Such filters which are used for the restoration of degraded images are usually called *deconvolution filters*. If we are successful in determining the restoration filter, then the restoration job is almost done.

The problem of finding the restoration filter depends on the type of degradation and noise that has affected the original image. The additive noise is generally random in nature and is position-independent. The degradation function is usually estimated using various techniques. The necessity for estimating the degradation function comes about because the actual degradation function is seldom known completely. One of the ways to estimate the function is to gather information from the image itself. A sub-image is constructed by using sample intensity values from the observed or the degraded image. The sample intensity values are picked up by looking into the background and the foreground of an object in the image where we feel there is a stronger signal content. By marking the pixel locations which are used to create the sub-image and picking out the observed image pixel values from the same locations, we can estimate the degradation function. Let us denote the sub-image of the observed image (in frequency domain) as $G_s(u, v)$ and the manually recreated sub-image as $\tilde{F}_s(u, v)$. The degradation function can be found out by

$$H(u, v) = \frac{G_s(u, v)}{\tilde{F}_s(u, v)}$$

The assumption here is that H is position invariant. Now, this can be used to reconstruct the total image from the full observed image, G. This manner of estimating the degradation function is known as estimation by image observation. Clearly, the problem with this method is that it cannot be made automatic in nature. The most popular way of estimating the degradation function is to mathematically model the degradation function. The Gaussian model is the most

popular of the models that is used for the estimation of the degradation function. The Gaussian filter is usually suitable to model mild and uniform blurring. Once the noise and the degradation functions have been estimated, then we need to use them to restore the original image. One of the popular restoration methods is to use inverse filtering. This is the simplest approach to restore the images once we have the degradation function and the knowledge of the type of noise. The process is performed in the frequency domain. Let us denote the reconstructed image as $\tilde{F}(u, v)$.

$$\tilde{F}(u, v) = \frac{G(u, v)}{H(u, v)}$$

The division of pixel by pixel values of the transform of the observed image and the degradation function will give the estimated pixel value of the reconstructed image. From Eq. (3.2), we know that,

$$G(u, v) = H(u, v) \cdot F(u, v) + N(u, v)$$

thus, substituting in the place of $G(u, v)$ in the previous equation we obtain Eq. (3.4).

$$\tilde{F}(u, v) = F(u, v) + \frac{N(u, v)}{H(u, v)} \tag{3.4}$$

The problem with the method is that we do not exactly know the Fourier representation of the noise term. Another problem being, the second term in the equation will tend to dominate if the values of $H(u, v)$ are too small or zero. To go around this problem, the Mean Square Error Filtering is used. This method is also called the *Wiener Filter* method [Gonzalez and Woods, 2003]. The wiener filter estimate focuses on solving the problem given in Eq. 3.5.

$$\text{Minimize } e^2 = E\{(f - \tilde{f})^2\} \tag{3.5}$$

where e is the mean square error between the original and the estimated image and $E\{.\}$ is the expected value. After solving Eq. 3.5 the following is obtained.

$$\begin{aligned}\tilde{F}(u,v) &= [\frac{H^*(u,v)S_f(u,v)}{S_f(u,v)|H(u,v)|^2 + S_\eta(u,v)}]G(u,v) \\ &= [\frac{1}{H(u,v)}\frac{|H(u,v)|^2}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}]G(u,v) \end{aligned} \tag{3.6}$$

where the terms in Eq. 3.6 are as follows:
$H(u,v)$ = Degradation Function
$H^*(u,v)$ = Complex conjugate of H
$S_\eta(u,v) = |N(u,v)|^2$ = Power spectrum of the noise
$S_f(u,v) = |F(u,v)|$ = Power spectrum of the undegraded image

The drawback with the method is that it requires information about the actual image. This information is usually not available. To accommodate this fact the method is slightly changed as shown in Eq. (3.7).

$$\tilde{F}(u,v) = [\frac{1}{H(u,v)}\frac{|H(u,v)|^2}{|H(u,v)|^2 + K}]G(u,v) \tag{3.7}$$

K is a parameter that represents the ratio of the noise and the actual image power spectrums. If the noise is assumed to be *white noise* then the power spectrum of $N(u,v)$ is constant. Then the only term that needs to be estimated is the power spectrum of the original image, i.e., $S_f(u,v)$. The choice of K is performed empirically. Wiener filter produces better results than the inverse filter technique. However, the problem of Wiener filter estimation is that it makes an assumption that the undegraded image and the noise belong to homogeneous random fields and that their power spectra are known [Rosenfeld and Kak, 1982]. In most of the situations one may not have *a-priori* knowledge to this extent.
Constrained Least Square filtering method is used if the mean and variance of noise is known. The method requires no priori information about the original image. The methods discussed till now assume that the degradation process is linear in nature. The linearity assumption works fine in most of the standard

restoration problems but tends to fail when there is nonlinearity associated in the image recording processing. The *Maximum a-Posteriori* (MAP) method is one such nonlinear image restoration process. The MAP method is capable of taking the nonlinearities encountered in the image recording process and also allows the ensemble mean of the image random process to be non-stationary. We will discuss more about the MAP technique in the next section.

3.1.2 Proposed Model for Demosaicking

The previous section gave an overview of the traditional image restoration model and the popular techniques used to restore the image from its degraded version. Continuing the discussion on modelling the demosaicking process as an image restoration process, we try to develop a mathematical model for the demosaicking process. For simplicity let us for now leave out the possibility of degradation and noise added to the image due to external influences. The demosaicking process basically deals with reconstructing a mosaicked image to a multi-spectral image. That is, the degradation that is being attempted to overcome is the absence of some of the pixels in the image plane. This is overcome by using interpolation algorithms in the traditional demosaicking methods. In the present scenario, this degradation has to be overcome by considering the demosaicking problem as an image restoration process. The mosaicked image, i.e. the image formed by the samples from the multi-spectral image by using a MSFA, can be considered as the degraded version of the actual image. So now we have identified the degraded image (the mosaicked image), the type of degradation (the missing pixels) and the required output (the final reconstructed multi-spectral image). This should be enough for us to form an image restoration type of model. Unfortunately, the modelling process is not a simple task in this case. The mosaicked image, which is the observed and the degraded image is a single channel image formed from systematic sampling of the original multi-spectral image. The degradation, which is the missing pixels in the image plane needs to be modelled mathematically so that it can be analyzed and included in the restoration model. The output needs to be a multi-spectral image close enough to the original image. There is no infor-

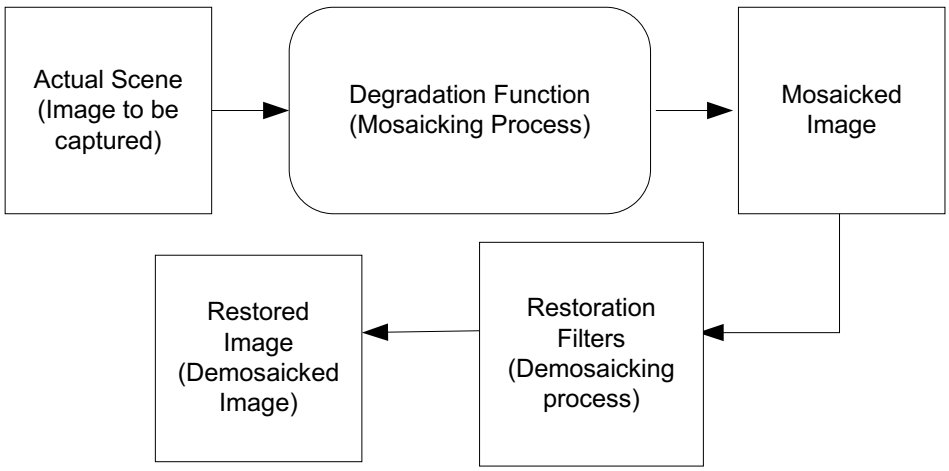


Figure 3.4: Demosaicking as Image Restoration Process

mation available of the original image, so no prior knowledge about the original image can be used in the modelling process. Moreover, the demosaicking process being an interpolation process which is position dependent cannot be modelled as a *convolution* process between the original image and the blur kernel (i.e., h [1] , refer to Sec. 3.1.1). The Fig. 3.4 gives a block diagram of the demosaicking process considered as a restoration problem. The block diagram is a simple description of how each process within the traditional image restoration problem can be made to relate to the demosaicking process. However, it must be kept in mind that the block diagram does not mean the exact mathematical model that was used for the restoration problem can be used for the demosaicking problem. Before we give a mathematical model for the demosaicking problem we have to look into another aspect of the demosaicking problem.

Now we come to the point where the external degradation and noise has to be taken into consideration. The modelling of these two quantities is similar to the tradition image restoration procedures. The degradation may occur due to various factors like defects of optical lens, non-linearity of electro-optical sensors, relative

[1]The blur kernel or the linear position-invariant spatial version of the degradation function, H, is also popularly known as the *point spread function* (PSF). The name comes from the fact that all the optical systems tend to blur a point of light to some extent. The amount of blurring being determined by the quality of the optical components. [Gonzalez and Woods, 2003]

Table 3.1: Comparison Between Image Restoration and Demosaicking Processes

.	**Demosaicking**	**Image Restoration**
g	Mosaicked Image	Degraded Image
f	Actual Multi-Spectral Scene	Original Image
Degradation	Missing pixel values	external noise and blur only
Output Image	Demosaicked Image	Reconstructed Image

motion between the camera and object, wrong focus, atmospheric turbulence, dust, etc. [Sonka, 2004]. For simpler understanding, the degradation may be assumed to be a linear position-invariant degradation. The noise model can be assumed as random additive noise which is also position invariant. However, when we consider the blur and noise for multi-spectral images, the degradations may be channel dependent. That is, the kind of blur and noise added to each of the spectral bands may be different and dependent on the spectral band that is being captured. All these factors have to be considered while modelling the demosaicking problem as an image restoration problem. The Table 3.1 gives an overview of the analogies made between the demosaicking process and the standard image restoration problem. The table gives an even more clear understanding towards the modelling process.

Mathematical Description of Model The image restoration model described by Eq. 3.1 can be modified to Eq. 3.8 to accommodate the non-linearity in the image formation process [Andrews and Hunt, 1977].

$$g(x,y) = s\{(f(x,y) * h(x,y)) + \eta(x,y)\} \tag{3.8}$$

The function s represents the mosaicking process and the notation $s(\mathbf{x})$ means that every element of the vector $\mathbf{x}$ is transformed by the function s. In the case of demosaicking, Eq. 3.8 may be translated as Eq. 3.9. This equation represents the mosaicking process of a degraded image.

$$m(x,y) = s\{(f(x,y,z) * h(x,y,z)) + \eta(x,y,z)\} \tag{3.9}$$

The original multi-spectral image $f(x, y, z)$ undergoes external channel dependent degradation $h(x, y, z)$. Additive noise $\eta(x, y, z)$ which also may be channel dependent is added to the degraded image. This forms a corrupted version of the original image. This degraded multi-spectral image is then given as an input to the mosaicking process. This results in the formation of the mosaicked image $m(x, y)$ which is a two-dimensional collection of the samples from all the spectral bands in the multi-spectral image. Taking the inverse of s on both sides of Eq.(3.9) we get

$$s^{-1}\{m(x, y)\} = s^{-1}\{s\{(f(x, y, z) * h(x, y, z)) + \eta(x, y, z)\}\}$$

The s^{-1} function actually denotes the demosaicking process. The left hand side of the equation means the demosaicking of the mosaicked image which is the demosaicked output. Let us denote the image by $g(x, y, z)$. Thus, the equation is translated into Eq. 3.10.

$$g(x, y, z) \approx (f(x, y, z) * h(x, y, z)) + \eta(x, y, z) \tag{3.10}$$

From Eq. 3.10, we see that the equation boils down to a multi-channel image restoration process. The model deals with one assumption, i.e., since the degradation caused due to missing pixel cannot be clearly modelled mathematically, a simple demosaicking is initially performed. This will fill the missing pixels but will leave out an optical blur due to the averaging operation from the interpolation process. This blur can then be accounted for by the $h(x, y, z)$. The assumption does not change the demosaicking problem to a great extent. The problem of finding the exact missing pixels and eliminating the external blur and noise still remains. Now that we have a mathematical model for the demosaicking process as a restoration problem, we need to find a way to reconstruct the estimate of the original multi-spectral image. The next section deals with possible ways to reconstruct the multi-spectral image from the degraded version.

3.2 MAP Technique for Restoration

Now that the mathematical model for the demosaicking problem is formed we need to use an appropriate method to perform the restoration process. The Maximum a-Posteriori (MAP) method is a nonlinear technique used for image restoration. The MAP method tries to find an estimate of an image such that it maximizes the *a-posteriori* probability $p(f|g)$. The method is based on Bayesian estimation process. The technique attempts to solve the following problem:

$$\text{Maximize}_{\ \mathbf{f}}\ p(\mathbf{f}|\mathbf{g})$$

where $p(\mathbf{f}|\mathbf{g})$ is the *a-posteriori* probability distribution. $\mathbf{f}$ is the original image and $\mathbf{g}$ is the observed image from the image restoration model discussed in the Sec. 3.1.2. From Bayes law, we know that

$$p(\mathbf{f}|\mathbf{g}) = \frac{p(\mathbf{g}|\mathbf{f})P(\mathbf{f})}{p(\mathbf{g})} \tag{3.11}$$

Considering Eq. (3.11), the restoration problem is transformed as in Eq. (3.12).

$$\text{Maximize}_{\ \mathbf{f}}\ \frac{p(\mathbf{g}|\mathbf{f})P(\mathbf{f})}{p(\mathbf{g})} \tag{3.12}$$

The $p(\mathbf{g}|\mathbf{f})$ term denotes the conditional probability of $\mathbf{g}$ given $\mathbf{f}$, i.e., given the original image $\mathbf{f}$, $p(\mathbf{g}|\mathbf{f})$ gives the probability distribution of the value of $\mathbf{g}$. The $p(\mathbf{g}|\mathbf{f})$ term is called the *sensor model* as it gives a description of the noisy or stochastic processes that relate the original unknown image $\mathbf{f}$ to the measured image $\mathbf{g}$. The next term, $P(\mathbf{f})$ is called the *prior model* [Qi, 1999]. The prior model is the *a-priori* probability of the unknown image $\mathbf{f}$. The $p(\mathbf{g})$ is given by

$$p(\mathbf{g}) = \int_{\mathbf{f}} p(\mathbf{g}|\mathbf{f})P(\mathbf{f})d\mathbf{f}$$

which turns out to be a constant independent of $\mathbf{f}$. Thus, the problem in Eq. (3.12) is transformed as Eq. (3.13).

$$\text{Maximize}_{\ \mathbf{f}}\ \mathbf{K}p(\mathbf{g}|\mathbf{f})P(\mathbf{f}) \tag{3.13}$$

where $\mathbf{K} = 1/p(\mathbf{g})$. To solve the Eq. (3.13) we need to have more information about the prior model and the sensor model. Once these terms are known then a optimization strategy can be used to solve the MAP problem.

3.2.1 Sensor Model

The sensor model is the conditional probability $p(\mathbf{g}|\mathbf{f})$ which means that if a value of $\mathbf{f}$ is given and fixed, then what would be the variation in $\mathbf{g}$. To find the conditional probability, we first have to know the relationship between $\mathbf{g}$ and $\mathbf{f}$. This will enable us to formulate the way $\mathbf{g}$ changes according to $\mathbf{f}$. Revisiting, the restoration model we have

$$\mathbf{g} = \mathbf{f} * \mathbf{h} + \eta$$

this means,

$$\eta = \mathbf{g} - \mathbf{f} * \mathbf{h}$$

i.e., if $\mathbf{f}$ is known and fixed, then the only parameter that affects $\mathbf{g}$ is the noise distribution η. This means the probability distribution of the conditional probability $p(\mathbf{g}|\mathbf{f})$ will be the same as the noise distribution. The noise occurs due to unknown external forces like dust, equipment faults, etc. Generally, such unknown noise is assumed to be independent Gaussian noise with zero mean, i.e., $\eta \sim N(0, \sigma^2)$; σ being the standard deviation of the normal distribution [Snyder et al., 2000]. The probability distribution function (PDF) of the noise term would then be

$$p(\eta) = K_1 \exp\{-\frac{1}{2}\eta^T[\phi_\eta]^{-1}\eta\}$$

where η represents the random variable for the noise distribution. The $[\phi_\eta]$ is the covariance matrix of the noise distribution and K_1 is a normalizing constant.

Given the noise distribution, the sensor model will be given as in Eq. (3.14).

$$p(\mathbf{g}|\mathbf{f}) = K_3 \exp\{-\frac{1}{2}(\mathbf{g} - (\mathbf{f} * \mathbf{h}))^T [\phi_\eta]^{-1} (\mathbf{g} - (\mathbf{f} * \mathbf{h}))\} \tag{3.14}$$

The constants in the equation are left unspecified but do not affect the maximization process of the *a-posteriori* probability. This will be seen further in the section as the solution for the maximization problem is being derived [Andrews and Hunt, 1977].

3.2.2 Prior Model

The prior model characterizes the original image. It provides the required prior knowledge about the original image in the form of a probability distribution function (PDF) or some other ensemble statistics. Since there is no information about the original image, it is impossible to characterize the prior model exactly according to the original image. In such a case, an accurate estimate of the prior model is formed. Hunt [Andrews and Hunt, 1977] has argued that an appropriate density function to use for the original undegraded images is the multivariate Gaussian. The multivariate Gaussian has a PDF given in Eq. (4.4).

$$P(\mathbf{f}) = K_2 \exp\{-\frac{1}{2}(\mathbf{f} - \mathbf{f}_m)^T [\phi_f]^{-1} (\mathbf{f} - \mathbf{f}_m)\} \tag{3.15}$$

The gaussian distribution is centered around the ensemble mean, $\mathbf{f}_m$, of $\mathbf{f}$ with a covariance matrix of $[\phi_f]$. The most significant point to note here is that $\mathbf{f}_m$ is not constrained to have all its components equal to the same constant. In other words, different components of $\mathbf{f}_m$ can have different ensemble means. By components, we mean different pixels in the original image can have different ensemble means associated with them. This means that the image, when considered as a random field, is now allowed to be nonhomogeneous (nonstationary) in nature. This kind of model is definitely closer to real-time situations than the usually used assumption of stationarity in the linear restoration methods. Hunt [Andrews and Hunt, 1977] gives an excellent example to illustrate this point. Consider a collection of images of full face photographs of different individuals. Assuming the photographs are taken in a manner such that the face is centered and there is a black background

around the object. It is clear that the ensemble average for each pixel will be different; it will be approximately zero outside the face area since the background is black and will vary from pixel to pixel within the face area. Similarly, another term of interest is the covariance matrix. The covariance matrix gives the amount of variation of each pixel value from its individual ensemble mean. In the case of multi-spectral images, the covariance matrix can also give an idea about the cross-correlation between individual bands of the multi-spectral image. The use of *Markov Random Fields* (MRF) for modelling the *a-priori* belief is also made in many applications. The MRF probabilistic model accounts for the local property of the images. Though, the MRF models the local property of the image accurately, it is very difficult to estimate the Markov distribution directly from the conditional distribution. Hammersley *et al.* have found an equivalence to the MRF model. The Gibbs distribution seems to have shared close equivalence to the MRF model with respect to modelling the local characteristics of the image through energies which describe the relationships shared by the neighboring pixels with each other. Eq. (3.16) gives a basic definition of the Gibbs distribution.

$$P(\mathbf{f}) = \frac{\exp(-U(\mathbf{f})/T)}{Z} \tag{3.16}$$

where the normalizing constant is$Z = \sum_{\mathbf{f}} \exp(-U(\mathbf{f})/T)$. Z is also called the *partition function*. The temperature of the model is denoted by T and $U(\mathbf{f})$ denotes the energy function. The energy function is given by

$$U(\mathbf{f}) = \sum_{i} V_{C_i}(\mathbf{f})$$

where C_i is the group of neighborhood pixels around the pixel i and $V_{C_i}(\mathbf{f})$ is called the *potential*. The potential depends on the local characteristics of the image. To find the potential at a particular pixel location in the image, we have to consider a set of pixels in the neighborhood of the pixel and find the difference between the pixel value and its neighbors. This is modelled in the form of a penalty function. The function is designed such that it penalizes the noise, but not the edges [Snyder et al., 2000]. It has been found that the inverted Gaussian function

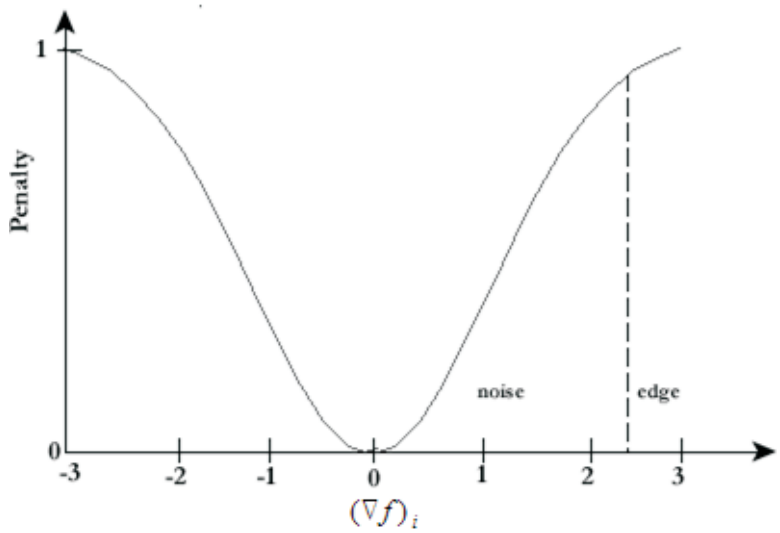


Figure 3.5: Inverted Gaussian as Penalty Function

produces noise removal without blurring the edges. The inverted Gaussian is shown in Fig. 3.5. The plot is between the amount of penalty to the difference between the intensity values of the center pixel and its neighbors. As we see from the figure, the noise in the image is penalized and at an occurrence of an edge the penalty is equal to 1, i.e., no penalty is imposed. Using this information, the prior term is modelled as Eq. (3.17).

$$P(\mathbf{f}) = Z^{-1} \exp\left(-\frac{\sum_i -\frac{\beta}{\sqrt{2\pi}\tau} \exp\left(-\frac{(f \otimes r)_i^2}{2\tau^2}\right)}{T}\right) \tag{3.17}$$

where the parameter β is used to adjust the smoothness of the image. The convolution term comes from the $\nabla \mathbf{f}_i$ term in the inverted gaussian distribution. The term signifies the difference between the center pixel and the neighbors of the center pixel. The difference is calculated by using a derivative kernel r. The *Laplacian operator*, a second derivative operator, may be used as the r kernel. Another operator that can be chosen as the r kernel is called the *Quadratic Variation*. The quadratic variation has a typical characteristic of never being negative which makes the edge more stable. Eqs. (3.18) and (3.19) show the Laplacian

and the quadratic variation kernels respectively.

$$r_L = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \tag{3.18}$$

The quadratic variation technique uses three kernels. They have been derived from the following formulation:

$$(\nabla f)^2_{i,j} = (\frac{\partial^2 f}{\partial x^2})^2 + (\frac{\partial^2 f}{\partial y^2})^2 + (\frac{\partial^2 f}{\partial x \partial y})^2$$

$$\begin{aligned} r_{xx} &= \frac{1}{\sqrt{6}} \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \\ r_{yy} &= \frac{1}{\sqrt{6}} \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\ r_{xy} &= \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \end{aligned} \tag{3.19}$$

3.2.3 Solving the MAP Problem

After getting thoroughly acquainted with the sensor model and the prior model, it is time to solve the optimization problem. Before we go into the mathematical formulation of the MAP solution, we need to understand that the images in this formulation are being considered in *lexicographic ordering* or *stacked notation*. This notation of images enables the consideration of the two-dimensional image plane as a vector. The *lexicographic ordering* notation has been proved to be computationally efficient [Andrews and Hunt, 1977] [Rosenfeld and Kak, 1982]. The use of this notation enables the most general linear system relating α and $\mathbf{g}$ to be given by $\mathbf{g} = [H]\alpha$. The traditional counterpart of the *stacked notation* is the

separable notation which looks at the same problem as $[G] = [U][\alpha][V]^T$. The earlier notation involves $[H]$ which has dimensions $N^2 \times N^2$ implying N^4 possible degrees of freedom (where N is the dimension of a square image). Whereas in the later case, the degrees of freedom available are onle $2N^2$. Therefore, the *stacked notation* offers a more general description of linear relationships between α and $\mathbf{g}$ than the *separable notation* case. We use the lexicographic ordering notation for images involved in the MAP problem. A uniform representation is used throughout this chapter for denoting the images in stacked notation. The image variables in bold letters are the images that are being considered in the *lexicographic ordering* notation.

Continuing with the solution of the MAP problem, we see that the MAP problem is a non-linear optimization problem with the objective function given in Eq. (3.13) as $\mathbf{K}p(\mathbf{g}|\mathbf{f})P(\mathbf{f})$. Substituting the expressions for the sensor model and the prior model (assume Gaussian case), we get

$$\text{Max}_{\ \mathbf{f}}\ \mathbf{K} \exp\{-\frac{1}{2}(\mathbf{g} - (\mathbf{f} * \mathbf{h}))^T [\phi_\eta]^{-1} (\mathbf{g} - (\mathbf{f} * \mathbf{h}))\} \exp\{-\frac{1}{2}(\mathbf{f} - \mathbf{f}_m)^T [\phi_f]^{-1} (\mathbf{f} - \mathbf{f}_m)\} \tag{3.20}$$

For ease of solution, the objective function is transformed into an easier form. However, we need to ensure that the change in the form of the objective function does not alter the problem. Taking log on both sides of the Eq. (3.13) we obtain a new objective function given in Eq. (3.21).

$$\ln p(\mathbf{f}|\mathbf{g}) = -\frac{1}{2}(\mathbf{g} - (\mathbf{f} * \mathbf{h}))[\phi_\eta]^{-1}(\mathbf{f} * \mathbf{h}))^T - \frac{1}{2}(\mathbf{f} - \mathbf{f}_m)[\phi_\mathbf{f}]^{-1}(\mathbf{f} - \mathbf{f}_m)^T - K \tag{3.21}$$

The log operation does not change the basic problem because logarithmic functions are increasing in nature. Now the left hand term in Eq. (3.20), $\ln p(\mathbf{f}|\mathbf{g})$ becomes the new objective function for the optimization problem. Let us denote the objective function by $W(\mathbf{f}) = \ln p(\mathbf{f}|\mathbf{g})$. Now, the function will be maximum

when $\nabla W(\mathbf{f}) = 0$. Thus,

$$\begin{aligned} \nabla W(\mathbf{f}) &= \nabla \ln p(\mathbf{f}|\mathbf{g}) \\ &= [H]^T[\phi_\eta]^{-1}(\mathbf{g} - [H]\mathbf{f}) - [\phi_\mathbf{f}]^{-1}(\mathbf{f} - \mathbf{f}_m) \\ &= 0 \end{aligned} \tag{3.22}$$

where $[H]$ represents the blur matrix. The spatial domain representation of the Eq. (3.22) is given as follows

$$\begin{aligned} \nabla W(\mathbf{f}) &= \nabla \ln p(\mathbf{f}|\mathbf{g}) \\ &= [\phi_\eta]^{-1}(\mathbf{g} - (\mathbf{f} * \mathbf{h})) * \mathbf{h}_{rev} - [\phi_\mathbf{f}]^{-1}(\mathbf{f} - \mathbf{f}_m) \\ &= 0 \end{aligned} \tag{3.23}$$

where $\mathbf{h}_{rev}$ is the reversed kernel of the spatial blur kernel, $\mathbf{h}$ [Qi, 1999]. If there exists a value of $\mathbf{f}$ such that the Eq. (3.23) is satisfied, then the value is called the *MAP estimate* and designated by the notation $\tilde{\mathbf{f}}_{MAP}$. Assuming $\tilde{\mathbf{f}}_{MAP}$ exists then rearranging Eq. (3.23) we get

$$\tilde{\mathbf{f}}_{MAP} = \mathbf{f}_m + [\phi_\mathbf{f}][H]^t[\phi_\eta]^{-1}(\mathbf{g} - [H]\tilde{\mathbf{f}}_{MAP}) \tag{3.24}$$

This is a non-linear equation in $\tilde{\mathbf{f}}_{MAP}$ and since $\tilde{\mathbf{f}}_{MAP}$ appears on both sides, there is a feedback structure. One way to obtain a solution for $\tilde{\mathbf{f}}_{MAP}$ is by trial-and-error method. However, this kind of method is practically not possible to perform. In such a case, optimization techniques are used to obtain the value of $\tilde{\mathbf{f}}_{MAP}$. In this thesis, we have used the *Gradient descent method* to obtain the solution for Eq. (3.24). Maximizing the *a-posteriori* density is equivalent to minimizing the objective function, $W(f)$ (refer to Eq. 3.20). The gradient descent method may be used to minimize $W(f)$ according to the following,

$$\mathbf{f}_{k+1} = \mathbf{f}_k - \alpha \nabla W(\mathbf{f}) \tag{3.25}$$

where k represents the number of iterations and α controls the speed of convergence of the process. The method has been proved to have a guaranteed con-

vergence. However, the convergence is limited only to local optimal points. The overall equation for the gradient descent method is given in Eq. (3.26).

$$\mathbf{f}_{k+1} = \mathbf{f}_k - \alpha\{[\phi_\eta]^{-1}(\mathbf{g} - (\mathbf{f} * \mathbf{h})) * \mathbf{h}_{rev} - [\phi_\mathbf{f}]^{-1}(\mathbf{f} - \mathbf{f}_m)\} \tag{3.26}$$

A solution for the MAP problem with Gibbs distribution as the prior term can be developed on similar lines as the previous case. The Eq. (3.27) gives the expression for this case.

$$\mathbf{f}_{k+1} = \mathbf{f}_k - \alpha[-[\phi_\eta]^{-1}(\mathbf{g} - (\mathbf{f} * \mathbf{h})) * \mathbf{h}_{rev} + \{\frac{\beta}{\sqrt{2\pi}\tau^3}(\mathbf{f} * \mathbf{r}) \exp(-\frac{(\mathbf{f} * \mathbf{r})^2}{2\tau^2})\} * \mathbf{r}_{rev}] \tag{3.27}$$

where $\mathbf{h}$ and $\mathbf{r}$ denote the blur kernel and the gradient kernel (Laplacian or quadratic variation) respectively. $\mathbf{h}_{rev}$ and $\mathbf{r}_{rev}$ are the flipped versions of the blur and gradient kernels respectively in the vertical and horizontal directions.

This completes the formulation of the demosaicking process as an image restoration problem. The problem is solved using the MAP technique using the gradient descent optimization process. In the experimental process, both the Gaussian and the Gibbs distributions were tested for reconstructing the multi-spectral images. The experimental results and detailed discussions related to the results are given in Chapter 4.

Chapter 4

Experimental Results and Discussions

However beautiful the strategy, you should occasionally look at the results.

– Winston Churchill

The previous chapters discussed about the details of the Mosaic Focal Plane Array (MFPA) technology that can be used for multi-spectral images. The methods have been tested and analyzed for their merits and demerits. In this chapter we will discuss the implementation and the results obtained from each of the strategies discussed earlier. Before we start looking into the results obtained from the experimentation process, it is important to look briefly into the way the experiments have been conducted.

4.1 Experimental Image Database

The experiments were conducted on seven band multi-spectral images. The images were provided by the group at North Carolina State University (NCSU). A set of eight multi-spectral images, each containing seven spectral bands were used for the experimental process. These images are synthetic images that have

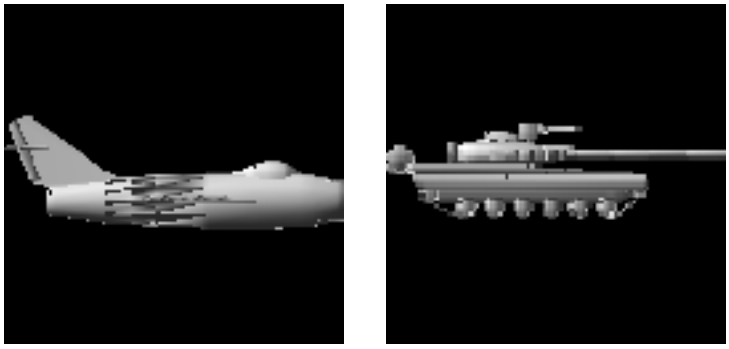
been created to have three visual (RGB) bands, three Mid-wavelength Infrared (MWIR) bands (3 - 8μm) and one Long-wavelength (LWIR) band (8 - 15μm). The synthetic images serve as a good basis for experimentation. The existence of seven bands from various spectral bands including the infrared spectrum makes the experimental process more appropriate to emulate actual applications. Most of the multi-spectral images are used in defense related applications, agriculture, etc., where the use of spectral bands other than the visible spectrum is quite common. For instance, in defense related applications night vision is obtained by the use of infrared spectral bands. Fig. 4.1 shows the set of eight multi-spectral bands used in the experiments. The figure displays the first band of each of the eight multi-spectral images. The methods developed in this thesis pertain to the development of a demosaicking methods for multi-spectral camera that use the MFPA technology. The best way of testing these methods would be to consider real-time multi-spectral images. But due to the unavailability of correct ground truth for real multi-spectral images, we have go for images that have been synthetically developed. These synthetic images were considered to be actual multi-spectral scenes for the purpose of simulating the experiments. The images have been created in a special image format called the *IFS* format. The *IFS* format, which stands for *Image File System*, has been widely used by the Image Analysis Laboratory, NCSU. This format has been specially designed to handle multi-spectral images, but at the same time it is capable of handling any type and size of images [Snyder, 1991].

4.2 Performance Metrics

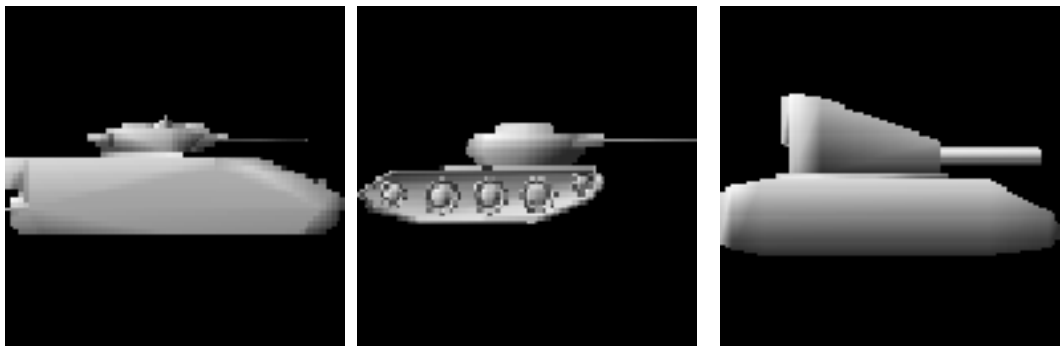
There is a need for a common platform that can be used to compare the results obtained from various methods implemented in this thesis. This platform will enable us to comment on the performance of each of the methods and choose the method that is best suitable for the camera. The design of the comparator depends on the application of the output of the multi-spectral camera. Generally, the multi-spectral cameras are used for classifying the objects being photographed. Unlike visual cameras, where the objective is to obtain the best visually appeal-



(a) 747 (b) Dc10 (c) F15



(d) Mig (e) Tank0



(f) Tank1 (g) Tank2 (h) Tank3

Figure 4.1: Database of Multi-Spectral Images

ing output, the multi-spectral output needs to have the best retained shape of the object-in-focus. Though creating the best visually appealing image does not seem very different than creating the best shape-retained image, they are still two different concepts altogether. An accurate shape-restored image may not be the best visually appealing image. This will become clearer as we go on to discuss the different metrics used for comparing images. The most basic method for comparing images is through visual evaluation. This kind of comparison works only in the case of visual images and tends to fail if the difference between the two images is too subtle for the human eye to observe. In such cases, there is a need for comparators which perform the comparison based on statistical characteristics of the images. Two such statistical methods have been used to compare the multi-spectral demosaicking outputs.

4.2.1 Reconstruction Accuracy

This metric calculates the amount of closeness between two images. The reconstruction accuracy is measured by calculating the Mean Square Error (MSE) between the two images that are being compared. If P and Q are two multi-spectral images, then the MSE calculation between P and Q is given as in eq. (4.1).

$$(MSE)_{P,Q} = \sum_{k=1}^{Bands} \sum_{i=1}^{Rows} \sum_{j=1}^{Columns} (P(i,j,k) - Q(i,j,k))^2 \tag{4.1}$$

Imagine $P(i,j,k)$ and $Q(i,j,k)$ are two points in the three-dimensional space, then from the Eq. (4.1) we can see that the MSE actually gives the distance (square of distance, to be more precise) between the two points. The MSE value is lesser implies the two points are close to each other. Therefore, looking at the overall images, if P and Q are similar to each other, the distance between them is lesser than the case where P and Q are different than each other. That is, the lesser the MSE value, the closer or the more similar the images are to each other.

This is used as one of the performance measures to evaluate the performance of the demosaicking methods. The objective of the demosaicking method is to create a multi-spectral output that is as close as possible to the original image.

The original image in our experiments are available as synthetic images listed in Fig. 4.1. The reconstruction accuracy measures the accuracy with which a demosaicking algorithm is capable of reconstructing the original scene. That is, this metric is capable on commenting on the quality of the reconstruction process of the demosaicking method. An example of the use of MSE as a performance measure for demosaicking can be found in Table 1.1 in Chapter 1. The table gives the $MSEs$ of different RGB demosaicking methods.

4.2.2 Classification Accuracy Using Spatial Information

The measure of the extent to which the shape of an object in an image has been retained is given by the spatial accuracy metric. The metric consists of shape statistics which can be used to characterize the shape of an object in an image. We use the $Hu - Moments$ for characterizing the shape of the object of interest. The *Hu Moments* are a set of seven numbers that describe the shape of an object in the image. These statistical values have been developed to be invariant to translation, rotation and scaling of the object of interest. For this reason, the *Hu Moments* are also popularly known as *Invariant Moments*. The following set of equations explain the procedure to calculate the invariant moments for an image. Consider an two dimensional image plane, f, then the moment of order $(p+q)$ is defined as Eq. (4.2).

$$m_{pq} = \sum_{x=1}^{Rows} \sum_{y=1}^{Columns} x^p y^q f(x,y) \tag{4.2}$$

Based on the moments calculation, the *central moments*are given as in Eq. (4.3).

$$\mu_{pq} = \sum_{x=1}^{Rows} \sum_{y=1}^{Columns} (x - m_h)^p (y - m_v)^q f(x,y) \tag{4.3}$$

where $m_h = \frac{m_{10}}{m_{00}}$ and $m_v = \frac{m_{01}}{m_{00}}$ represent the center of gravity about the horizontal and the vertical directions. The *normalized moments* can be easily calculate from

the central moments. The normalized moments are given by the following,

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}$$

where $\gamma = \frac{p+q}{2} + 1$. The second and third order normalized moments are then used to derive the set of invariant moments. The following set of equations gives the set of seven invariant moments that are used to characterize the shape of an object in the image.

$\phi_1 = \eta_{20} + \eta_{02}$
$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$
$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$
$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$
$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$
$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$
$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

Fig. 4.2 gives a general block diagram for extracting the invariant moments from an image. The input image undergoes segmentation so that the object-of-interest (foreground) and the background are separated. Image thresholding operation is used to segment the image. The invariant moments are then extracted from the segmented image by using the above given set of equations.

4.2.3 Classification Accuracy Using Spectral Information

The average intensity levels of each of the spectral bands in the multi-spectral image are used as the spectral features for calculating the spectral classification accuracy of the demosaicked outputs. In the context of this thesis work, we are using seven-band multi-spectral images. Thus, we will have a set of seven spectral features for calculating the spectral classification accuracy (i.e., one feature from each spectral band).

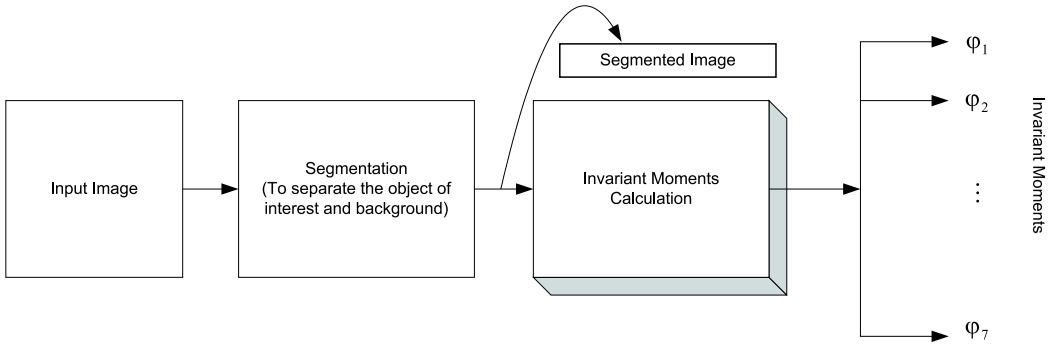


Figure 4.2: Block Diagram for Invariant Moments Calculation

4.2.4 Calculation of Classification Accuracy

One of the major objectives of using the MFPA technology in multi-spectral cameras is to improve the classification accuracy of the multi-spectral image acquisition system. The ability to classify or differentiate between different objects is measured using the classification accuracy. Classification accuracy plays a very important in varied applications like defense, agriculture, medicine etc where the main motive of obtaining a multi-spectral image is to successfully identify the object being photographed. This is done by collecting various images and training the network to identify each of the objects. In our case, we test for the classification accuracy based on the spectral and spatial characteristics of the multi-spectral database. The output from each of the methods is tested for its ability to successfully classify the object-of-interest. The generalized block diagram for performing the classification accuracy calculation is given in fig. 4.3. The classification process needs a training and a testing database. The training dataset is formed by scaling the object in the image and rotating it for 18 rotations and then extracting the spectral and spatial features by using the metrics mentioned earlier. The testing dataset is also formed in a similar manner, except that in case of the testing dataset the object is rotated for 20 rotations. The *K-nearest neighbors* (kNN) classifier is used for classifying the testing images based on the training dataset. The kNN classifier is a basic classifier which classifies the images into particular classes based on their distance from each of the elements in the training dataset.



Figure 4.3: Calculation of Classification Accuracy

The distances are then sorted in a descending and the object is assigned to the class based on the majority of classes present amongst the first K sorted distances. The value of K depends on the number of training data samples. This process is followed for all the images present in the database and the classification accuracy is calculated as the ratio between the number of correctly classified images to the total number of testing samples. The higher the classification accuracy, the better is the performance of the method. The classification accuracy is calculated with respect to the spectral and spatial accuracies of the images. As mentioned earlier, the spectral accuracy represents the capacity of the method to spectrally reconstruct the actual/ original image as closely as possible. And the spatial accuracy signifies the capacity of the method to retain the maximum shape information of the object being photographed.

4.3 Demosaicking Results For Multi-Spectral Images

This section will discuss the demosaicking results obtained on applying the MFPA technology to multi-spectral images. The experimental process will cover the detailed results for mosaicking and demosaicking processes discussed in Chapter 2. The experimental process is illustrated in Fig. 4.4. The original image is the synthetic image from the available seven band IFS image database. The original image goes through mosaicking to produce a single-band mosaicked image which is a systematic collection of all the seven spectral bands distributed throughout the image according to the seven band MSFA (see Fig. 2.7). This mosaicked image is sent as an input to the demosaicking block which tries to reconstruct the original multi-spectral image using various interpolation-based strategies discussed in the previous chapters. The demosaicked output is then tested for its performance by using the performance metrics discussed earlier. Fig. 4.5 shows the set of seven bands of the original multi-spectral image. This multi-spectral image is sent to the mosaicking block. The seven band mosaic filter array is applied on the original multi-spectral image resulting in the mosaicked image displayed in Fig. 4.6.

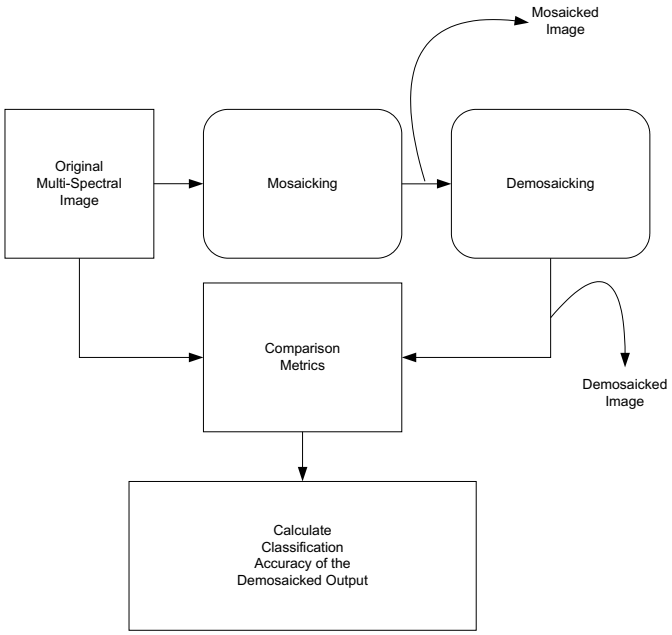


Figure 4.4: Experimental Process



Figure 4.5: Seven Bands of the Original Multi-Spectral Image

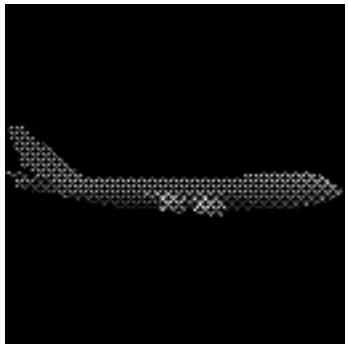


Figure 4.6: The Mosaicked Image



(a) Original Image

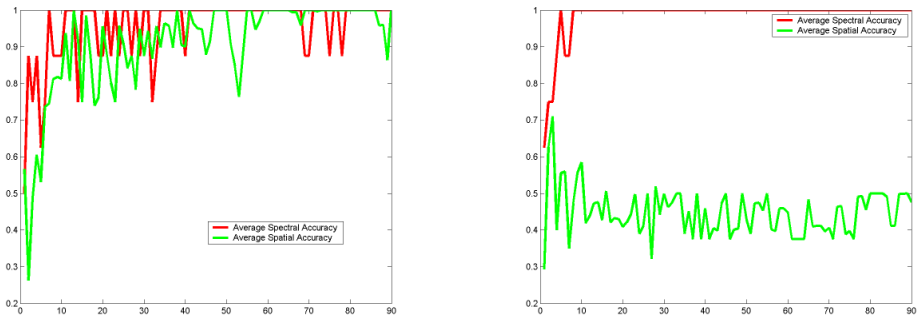


(b) Bilinear Demosaicked Image

Figure 4.7: Bilinear Demosaicking Results

4.3.1 Bilinear Demosaicking

The mosaicked image from Fig. 4.6 is demosaicked using the basic bilinear demosaicking algorithm. Fig. 4.7 displays the seven bands of the original and the bilinear demosaicked image. For better visual comparison, the images have been placed in the particular order in the figure. By visual evaluation it can be seen that the bilinear demosaicking result is not perfect and suffers from edge artifacts. The *zipper effect* can be clearly seen at edge locations in the output image compared to the original image. The bilinear demosaicking method does not involve any edge considerations and thus results in the formation of step-like edge artifacts. The classification accuracy curves for the bilinear demosaicking output have been given in Fig. 4.8. The classification curve in Fig. 4.8(a) has been derived from the set of original synthetic images in our multi-spectral database. This curve represents the ideal spectral and spatial accuracies where the output image is exactly similar to the original image. That is, it is assumed that the demosaicking process produces an output which is an exact replica of the actual image. This original classification curve will act as a comparison ground for the rest of the classification curves. The classification curve for the bilinear demosaicked output, as seen from Fig. 4.8(b), is unexpectedly better than expected. The spectral accuracy curve (red curve) almost follows in similar lines as the original spectral accuracy curve. This means the bilinear demosaicking process produces a result

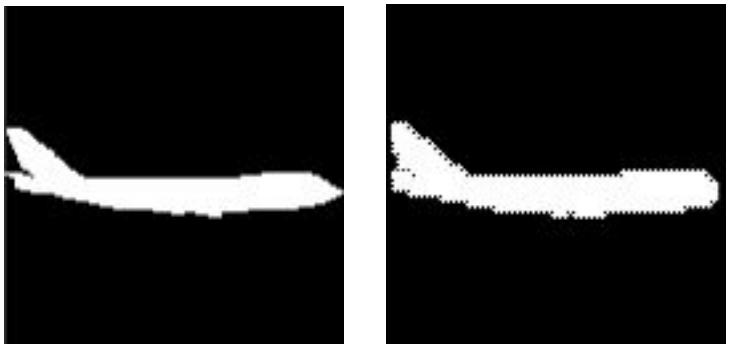


(a) Ideal Classification Curves

(b) Bilinear Demosaicked Classification Curves

Figure 4.8: Classification Accuracy Curves for Bilinear Demosaicked Output

that has similar intensity values as the original image. The result seems to be almost perfectly reconstructed spectrally. Although, the demosaicked output possesses a good spectral accuracy curve, the spatial response (green curve) seems to be badly affected. This is expected due to the inherent drawbacks of the bilinear demosaicking process of producing step-like artifacts at edge locations. The spatial classification curve is generated by considering the shape statistics of the most prominent band of the multi-spectral image. The most prominent band in the multi-spectral image is the one that has the highest sampling frequency and thus occurs the most number of times in the MSFA. The choice of the most prominent band for calculating the shape statistics is justified because being the highest sampled band makes the band have the most number of pixels in the mosaicked image which in turn means that the most prominent band has a greater amount of information pertaining to the shape of the object being photographed than in any other spectral band. In the case of the seven band multi-spectral images, the first band is the highest sampled band and thus is the one which is considered for extracting the shape statistics for the spatial accuracy curve. The *zipper effect* can be clearly observed in the comparison of the silhouettes of the original and the bilinear demosaicked outputs. Fig. 4.9 shows the comparison between the two silhouettes. The silhouettes are from the first band from each of the images. As we can see from the figure, the silhouette of the bilinear demosaicked output is clearly



(a) Silhouette of Original Image

(b) Silhouette of Bilinear Demosaicked Image

Figure 4.9: Zipper Effect in Bilinear Demosaicked Output

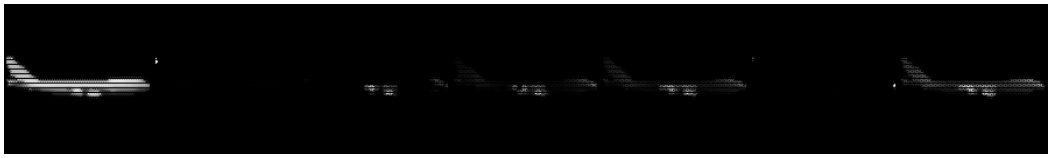
degraded and the degradation appears to be a step-like effect along the edges in the image. As we saw in chapter 2, the *zipper effect* in bilinear demosaicking is accounted for by using the Cok's demosaicking method.

4.3.2 Cok's Demosaicking

The Cok's demosaicking method has been inspired from the constant-hue based demosaicking method developed by Cok [Cok, 1987]. The demosaicking method performs excellently in the case of color images and has been suitably changed to be applied to the multi-spectral images. It still remains to be seen if the extension of the color based method is valid in case of the multi-spectral images. Fig. 4.10 lists the results obtained for the Cok demosaicking method. We see from the figure, that the results obtained from the Cok's demosaicking process are worse than the bilinear demosaicking results. This is unexpected because in color images, the Cok's demosaicking method works far better than the bilinear demosaicking method and is successfully used as a remedy for the step-like artifacts that creep into the bilinear demosaicking process. However, this result provides a perfect explanation for not extending the visual demosaicking processes directly to the multi-spectral case. The Cok's demosaicking method for visual images focuses on controlling the *hue* changes along the edges in the image. Although, *hue* is a property which is specific to color images only, we tried to give a logical extension



(a) Original Image



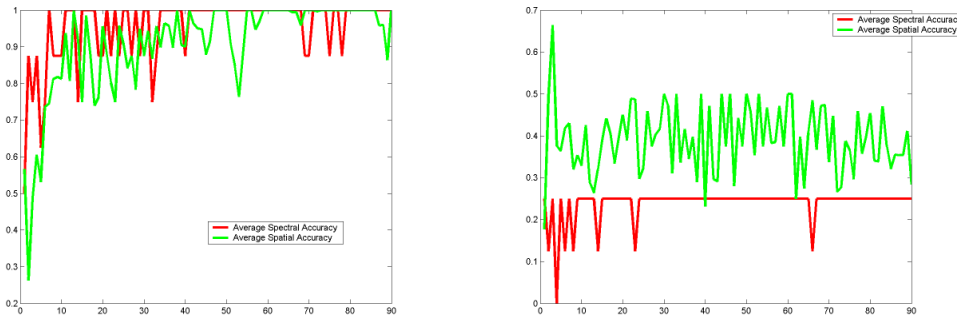
(b) Cok's Demosaicked Image

Figure 4.10: Cok's Demosaicking Results

to the method for the multi-spectral case by considering the most sampled band in the multi-spectral to play the same role as the luminance band in color images. From the results we see that this extension did not work as expected. This is because in color images the luminance and the two chrominance bands are correlated with each other. Whereas in the multi-spectral image, there is no such correlation. There is a possibility for some kind of correlation, but the relationship is not the same as the one shared by the spectral bands in a color image. Therefore, the use of the most prominent band in the multi-spectral image to interpolate the other bands cannot be justified. This fact is clearly being portrayed by the classification curves shown in Fig. 4.11. As we see from the figure, the spectral accuracy curve is severely affected. The spectral classification is worse than the bilinear demosaicking case. The spatial classification curve is almost similar to the bilinear demosaicking result. However, the classification curves do not serve any special purpose since the method does not seem appropriate to be used for multi-spectral images.

4.3.3 Median Based Demosaicking

The median based demosaicking method deals with a similar concept as the Cok's demosaicking method. The interpolation of missing pixels in such a manner such that the edge information in the image is preserved to the maximum extent. Like



(a) Ideal Classification Curves

(b) Cok's Demosaicked Classification Curves

Figure 4.11: Classification Accuracy Curves For Cok's Demosaicked Output

the Cok's method, this method performs the interpolation of missing spectral bands based on information from the most prominent spectral band in the multi-spectral image. Looking at the results from the Cok's method, we can predict that the median based demosaicking method will also fail to produce a good demosaicked output. The results have been listed in Fig. 4.12. Due to the median based process which smooths out edge artifacts in the demosaicked output by applying the median filter on difference images, the results from all the seven bands look almost similar. The main reason for such inappropriate result is because the seven bands in the multi-spectral image share different edge information. In case of visual images, the three spectral bands share almost similar edge information and that is why the median filtering of the difference images works perfectly in removing edge artifacts. But in the case of the seven band multi-spectral image, the edge information shared by all the bands is totally different from each other. One simple reason for the variation in edge information of the same object through the multi-spectral image is because of the nature of the spectral bands in the image. The multi-spectral image contains three RGB, three SWIR and one LWIR spectral band. The object projects itself in a different manner for each spectral band. This results in the variation of edge information throughout the multi-spectral image. For more clarity, see Fig. 4.13 for the difference in edge information in the seven bands of the original multi-spectral image. For the sake of completion,



(a) Original Image



(b) Median Based Demosaicked Image

Figure 4.12: Median Based Demosaicking Results

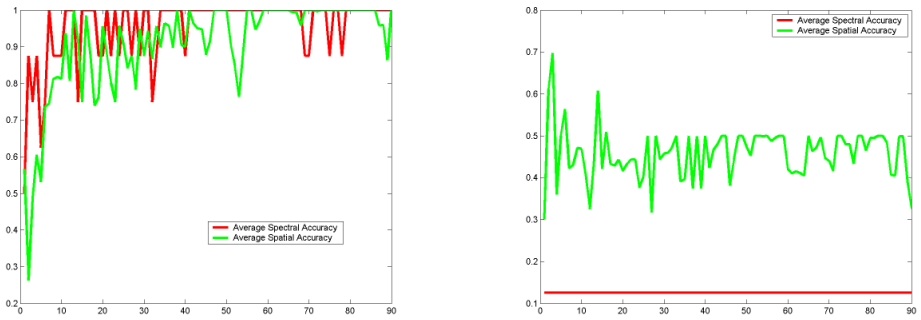


Figure 4.13: Variation of Edge Information in the Multi-Spectral Image

the classification curves for the median based demosaicking method have also been derived. The classification takes a serious hit in the spectral department. The spectral accuracy is almost nil compared to the original classification curves in Fig. 4.14(a). The spatial accuracy curve seems to be almost similar to the bilinear demosaicking classification result. The possible explanation for this is that, the spatial classification curve is generated only based on information from the first band in the multi-spectral image. The first band (or the most prominent band) seems to have been recreated better than the rest of the bands and thus provides a better spatial classification accuracy. However, due to the inherent drawbacks of the method the classification curves do not seem to signify anything useful.

4.3.4 Modified Bilinear Demosaicking

The modified bilinear demosaicking method has been derived as a reverse process of the seven band mosaicking process (see chapter 2 for details). This implies



(a) Ideal Classification Curves

(b) Median Based Demosaicked Classification Curves

Figure 4.14: Classification Curves for Median Based Demosaicked Output

that the method is perfectly suitable for the seven band multi-spectral images and should perform with great results. The results have been listed in Fig. 4.15. As predicted, the results are far better than the Cok's demosaicking method and the median based demosaicking method. Although, the demosaicking process seems to be designed perfectly still we can observed some artifacts at the edge locations. This is probably due to the averaging operation performed during the interpolation process. This is the same averaging operation which results in the step-like edge artifacts in the normal bilinear demosaicking process. The classification curves are listed in Fig. 4.16. The figure has three classification curves - the ideal classification curve, the bilinear demosaicking classification curve and the modified bilinear classification curve.

We observe from Fig. 4.16(b) the spectral classification accuracy of the modified demosaicked result is almost similar to the ideal classification accuracy. This means that the method is successful in producing a better reconstruction of the original multi-spectral image than the previously discussed methods. However, the spatial accuracy is not as better as the its spectral counterpart. The spatial accuracy curve shows an improvement over the normal bilinear demosaicking classification curve shown in Fig. 4.16(c). This makes the modified bilinear demosaicking method a better method than the other demosaicking methods. However, the problem that still persists is the improvement of the spatial classification accu-

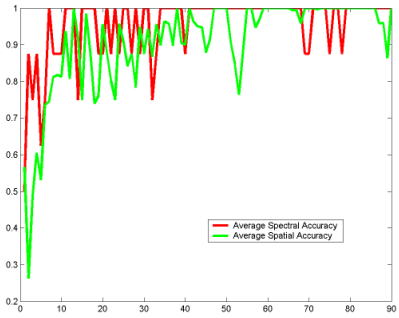


(a) Original Image

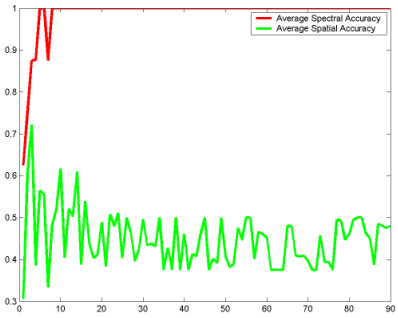


(b) Modified Bilinear Demosaicked Image

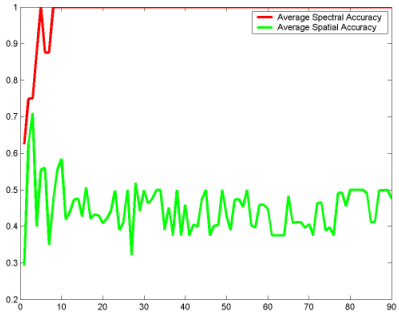
Figure 4.15: Modified Bilinear Demosaicking Results



(a) Ideal Classification Curves



(b) Modified Bilinear Demosaicked Classification Curves



(c) Normal Bilinear Demosaicked Classification Curves

Figure 4.16: Classification Curves for Modified Bilinear Demosaicked Output



(a) Original Image



(b) Modified Median Based Demosaicked Image

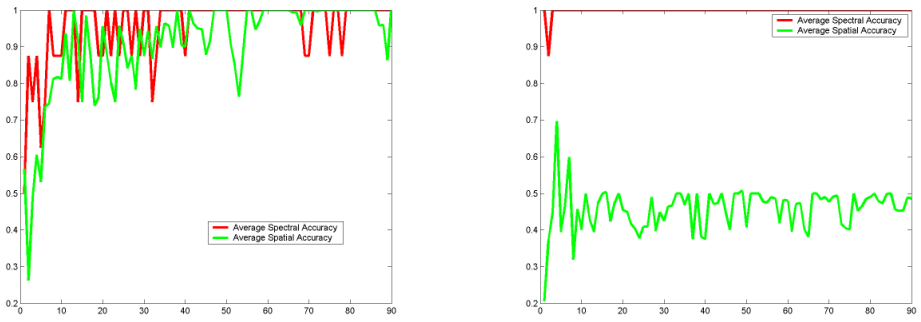
Figure 4.17: Modified Median Based Demosaicking Results

racy.The modified median based demosaicking method focuses on improving the output of the modified bilinear demosaicking result.

4.3.5 Modified Median Based Demosaicking

This method is a simple extension to the modified bilinear demosaicking method. The modified bilinear demosaicked output is passed through a median filter to remove the edge artifacts in the bilinear output. The focus of the method is to improve the spatial classification accuracy of the output image while keeping the spectral accuracy of the modified bilinear demosaicked output intact. The results obtained for this method are given in Fig. 4.17. We observe that the results do not differ much visually compared to the modified bilinear demosaicking output. The step-like edge artifacts have been reduced to a great extent. This means that we can expect an improvement in the spatial classification accuracy curves. Fig. 4.18 shows the comparison of the classification curves for the ideal and the modified median based case. The spatial accuracy has been improved compared to the modified bilinear demosaicking output.

Further analysis about the demosaicking results is performed by calculating the root mean square error (RMSE) with respect to the original image. The mean square error gives an idea about the ability of the demosaicking method to create a perfectly reconstructed image. The $RMSE$ values for all the demosaick-



(a) Ideal Classification Curves

(b) Modified Median Based Demosaicked Classification Curves

Figure 4.18: Classification Curves for Modified Median Based Demosaicked Output

Table 4.1: RMSE Values for Different Demosaicking Techniques

Demosaicking Method	**MSE Value**
Bilinear Method	8.7114
Cok's Method	14.2243
Median Based Method	12.3643
Modified Bilinear Method	5.3797
Modified Median Based Method	**4.9920**

ing methods are given in Table 4.1.We observe that the Modified Median Based Method gives the best reconstruction accuracy compared to the rest of the methods. The above table also shed some interesting facts about each of the methods. The Cok's method produces an output that has the worst $RMSE$ value compared to any of the other methods. This means the output of this method should have the least spectral accuracy compared to the rest of the methods. This has been shown to be right in Sec. 4.3.2. The $RMSE$ value improves in the median based demosaicking method, although the value still remains lesser compared to the bilinear demosaicking method. The bilinear demosaicking method being the most basic interpolation method with any considerations of edge-information or neighborhood sensitivities is supposed to produce the highest $RMSE$ value. Any other method which has a greater $RMSE$ value than the bilinear demosaicking method is bound to have major defects in it. The modified bilinear method pro-

duces an almost perfect reconstructed result. The ability of the process to revert the mosaicking process makes the reconstruction process to be near-perfect.

4.4 MAP Based Demosaicking Results

The use of MAP as an image restoration technique to demosaic multi-spectral images has been discussed at length in Chapter 3. In this section, we will discuss about the experiments conducted and the results obtained by using the MAP technique. The experimental process has been illustrated in Fig. 4.19. The original image is corrupted with noise and Gaussian blur. This degraded image is then mosaicked using the seven band MSFA. Then the MAP technique is used to recreate an undegraded version of the original image. The MAP process requires an initial estimate of the original image so as to start the iteration process. The Wiener estimate is used for this purpose. The experimentation process for the MAP technique has been performed for different prior and noise models. Revisiting the equation for the solution to the MAP problem when we use the Gaussian prior model, we have

$$\mathbf{f}_{k+1} = \mathbf{f}_k - \alpha[[\phi_\eta]^{-1}(\mathbf{g} - (\mathbf{f}_k * h)) * h_{rev} - [\phi_f]^{-1}(\mathbf{f}_k - \mathbf{f}_m)] \tag{4.4}$$

where the initial estimate of the original image is the value of $\mathbf{f}_k$ at $k = 0$. We use the Wiener estimate as the initial estimate. Though the wiener estimate does not suit the demosaicking problem, still it is better to have some kind of estimate than to start with a totally unknown estimate. This reduces the time of iterations and also increases the guarantee for a result. The initial estimate plays a very important role in the gradient descent procedure. The convergence to a local minimum through the gradient descent process totally depends on the initial starting point which is the initial estimate. From Eq. (4.4), h is the blur kernel. We choose the 3×3 Gaussian kernel as the blur kernel for the MAP process. h_{rev} is the reverse ordered version of h. In out case, since we assume h to be the gaussian blur kernel which is symmetric in both directions, $h_{rev} = h$. The noise term is an important consideration in the experimentation process. We test the

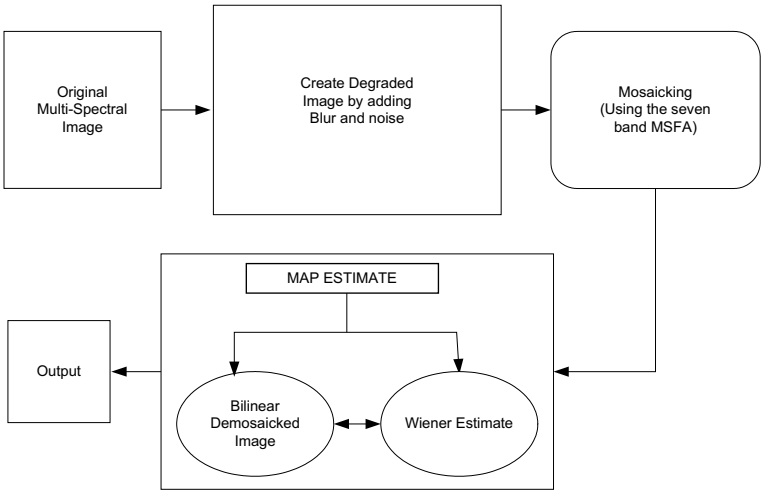


Figure 4.19: MAP Experimentation Process

system with variable variances and zero-mean gaussian noise term. $\mathbf{f}_m$ denotes the mean of the ensemble of the prior image set. The ensemble mean is unknown because we do not know the values of the images in the ensemble. However, we know that the distribution of the ensemble of prior images is Gaussian in nature. So as we go through each iteration, we start finding the values of the images in the ensemble of prior images. Then the $\mathbf{f}_m$ is the mean of all the prior images which have a gaussian distribution. The characteristics of the prior model are chosen to be constant throughout the process. The speed of convergence is controlled by the coefficient α. The process is considered to be converged when the difference between the $\mathbf{f}_{k+1}$ and $\mathbf{f}_k$ goes below some particular threshold value.

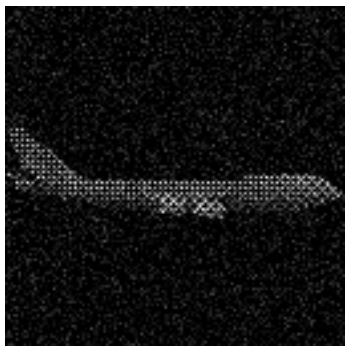
The original image (fig 4.20(a)) is corrupted by a gaussian noise and a gaussian distribution. The noise and distribution is same for all the seven bands. This gives us the degraded multi-spectral image, shown in fig 4.20(b). Next, the degraded image is mosaicked using the seven band MSFA to generate the mosaicked image (fig 4.20(c)). Comparing the original mosaicked image in Fig. 4.6 and the degraded version of the mosaicked image in fig 4.20(c). This image is the only information we have about the original image. If interpolation-based demosaicking methods are used to retrieve the multi-spectral image from the mosaicked image, it is guaranteed that we will end up with a totally degraded version of the multi-



(a) Original Multi-Spectral Image



(b) Degraded Multi-Spectral Image



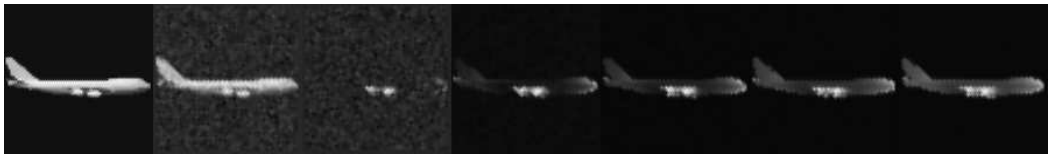
(c) Degraded Mosaicked Image

Figure 4.20: Original Image has been Degraded with Gaussian Noise and Blur. The Degraded Image is then Mosaicked to form the Degraded Mosaicked Image.

spectral image. Therefore, we adopt the MAP based algorithm to retrieve a proper estimate of the multi-spectral image. The Fig. 4.21 displays the results obtained after implementing the MAP based algorithm. We see that the method successfully retrieves a undegraded multi-spectral image. From the MAP result we observe some degradation still remaining in the some of the bands of the multi-spectral image. This happens due to two causes, firstly, the MAP based algorithm does not perfectly denoise and deblur the degraded image. Secondly, due to very low mean values of some of the bands in the multi-spectral image, any small intensity level of a pixel shows up onto the display. The best way to check whether or not the result has improved over the degraded image is to



(a) Degraded Image



(b) MAP Estimate Image (Gaussian Prior Model)

Figure 4.21: MAP Estimate Results Using Gaussian Prior Model

Table 4.2: Comparison of Degraded Image and MAP Estimated Image (Gaussian Prior Model)

Image	**MSE**	**PSNR (in dB)**
Degraded Image	45.0	31.5987
Bilinear Demosaicked Image	42.7655	31.81
MAP estimated Image	**40.078**	**32.1017**

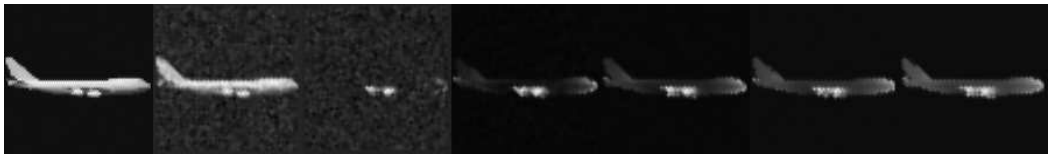
check for the *Mean Square Error*(MSE) values of the two images with respect to the original image. Another popular metric that is commonly used to compare image restoration outputs is the *Peak Signal-to-Noise Ratio*($PSNR$). The $PSNR$ metric gives an estimate of the amount of noise present compared to the actual signal strength in the image. Usually $PSNR$ has *decibels* as its units. The $PSNR$ calculation in dB iss given in Eq. (4.5).

$$PSNR_{dB} = 20\log_{10}(\frac{255}{\sqrt{MSE}}) \tag{4.5}$$

where MSE is the *mean square error* between the original and the existing image. The higher the $PSNR$ the better is the result. Table 4.2 gives the comparison between the degraded image and the restored image. We see from the table, that the MAP estimated image has a lesser MSE and a higher $PSNR$ than the bilinear demosaicked image. This means the MAP technique has successfully restored the seven-band multi-spectral image and performs better than the usual interpolation



(a) Degraded Image



(b) MAP Estimate Image (Gibbs Prior Model)

Figure 4.22: MAP Estimate Results Using Gibbs Prior Model (Laplacian Kernel)

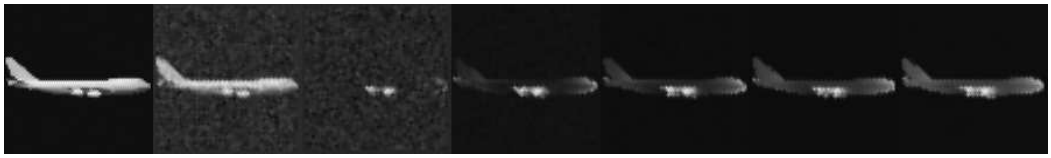
based methods in the presence of noise and other degradations.

Another prior model, the Gibbs distribution was also used to perform the MAP estimation. Revisiting, the discussion in Sec. 3.2.2 we see that the Gibbs distribution is supposed to model the image restoration problem better than the Gaussian model. The Gibbs prior model uses a difference calculation filter. Two such filters, *Laplacian* and *Quadratic Variation*, were used for performing the MAP estimation process. The results for the MAP estimate using the Laplacian and the Quadratic kernels have been listed in Fig. 4.22 and Fig. 4.23 respectively. Table 4.3 lists the MSE and $PSNR$ values for the MAP estimate with Gibbs prior model using the Laplacian kernel. The MSE value has been reduced which means the restored image is closer to the original image than the degraded image. The results for the Quadratic variation based Gibbs distribution model are listed in table 4.4. In this case also, the MSE values have been reduced but not as much as the Laplacian kernel based Gibbs distribution method.

The results listed in this chapter strengthens our claim for the use of MFPA technology for multi-spectral images. We have successfully tested two kinds of demosaicking strategies - the interpolation based and the image restoration based. The interpolation-based demosaicking methods work well in the absence of noise and degradations. For practical applications, where the absence of noise and degradations is almost impossible, we have developed the MAP based image restoration technique. Experiments were performed to test the performance of



(a) Degraded Image



(b) MAP Estimate Image (Gibbs Prior Model)

Figure 4.23: MAP Estimate Results Using Gaussian Prior Model (Quadratic Variation)

Table 4.3: Comparison of Degraded Image and MAP Estimated Image (Gibbs Prior Model with Laplacian kernel)

Image	**MSE**	**PSNR (in dB)**
Degraded Image	45.0	31.5987
Bilinear Demosaicked Image	42.7655	31.81
MAP estimated Image (Laplacian Kernel)	41.6734	31.9322

Table 4.4: Comparison of Degraded Image and MAP Estimated Image (Gibbs Prior Model Using Quadratic Variation)

Image	**MSE**	**PSNR (in dB)**
Degraded Image	45.0	31.5987
Bilinear Demosaicked Image	42.7655	31.81
MAP estimated Image (Quadratic Variation)	42.173	31.8804

the demosaicked images in classification based applications. The results have shown that the demosaicked images provide almost perfect spectral classification and about 70% spatial classification accuracy.

Chapter 5

Conclusions and Future Work

He did not arrive at this conclusion by the decent process of quiet, logical deduction, nor yet by the blinding flash of glorious intuition, but by the shoddy, untidy process halfway between the two by which one usually gets to know things.

– Margery Allingham

The use of MFPA technology has revolutionized the digital color camera industry. It has made it possible to develop smaller yet cheaper cameras which provide ample space for improving the image quality. The use of multi-spectral images has also risen to great heights. However, the multi-spectral camera market is yet to catch on with the use of a cheaper and a robust technology.

5.1 Contributions of Thesis

This thesis work provided a possibility for the use of the MFPA technology for multi-spectral cameras. New demosaicking algorithms were developed which are generic in nature and can be adapted for any kind of multi-spectral images. For practical applications of the technology, the work also addressed the possibility of degradations and noise creeping into the image acquisition process. A detailed

MAP based algorithm was developed to denoise, and deblur the multi-spectral image with a possibility of performing demosaicking at the same time. The use of this algorithm makes it possible to generate an optimal demosaicking process that will be automatic in itself and provide the best demosaicking result based on the parameters of the system. The results show an improvement over the interpolation-based strategies and strengthens the claim for the use of the MFPA technology for multi-spectral images. The use of multi-spectral images for applications like target recognition has also been addressed. The results for the classification accuracies of the demosaicked multi-spectral images have proved that the use of the MFPA technology not only comes with its traditional advantages but also is suitable for classification applications of the multi-spectral images.

5.2 Future Work

These are some things which we would like to do to extend the work in this thesis. There is a possibility to improve the image restoration algorithm. Instead of tackling the problem of removing the degradation and simultaneously filling the missing pixels, we can treat it as a pure demosaicking problem. Consider an ideal case of performing demosaicking on the mosaicked image. The mosaicked image can be further perceived as a combination of multiple images (with the number of images equal to the number of spectral bands in the multi-spectral image) with each spectral band image having a combination of existing and missing pixels. The missing pixels may be considered to have been produced due to a degradation upon the original multi-spectral image. If this kind of degradation (which causes some pixels to vanish) can be mathematically found, then its inverse could be used to revert the degradation effect. In other words, the degradation function can be used to find the missing pixels in the mosaicked image and thus producing a perfectly demosaicked output. The advantage of using such a technique is that firstly, unlike interpolation based strategies there is no need of a prior knowledge about the mosaicking array. Secondly, the demosaicking process will become purely automatic and will provide greater control over the kind of output produced.

A hardware simulation of the MFPA technology for multi-spectral images will

give a greater insight on the real advantages of the technology. The hardware simulation will also act as a guide of the creation of real world multi-spectral cameras.

The use of hexagonal tessellations for the detector geometry should be explored in detail. The hexagonal tessellations have proved to have an upper hand over the traditional square tessellations with respect to lesser sampling rates, lesser processing time, etc.

Bibliography

Bibliography

[Andrews and Hunt, 1977] Andrews, H. and Hunt, B. (1977). *Digital Image Restoration*. Prentice-Hall Signal Processing Series. Prentice-Hall Inc., 1st edition.

[Audley, 1997] Audley, M. D. (1997). Charge-coupled devices. Technical report, Laboratory for High Energy Astrophysics, NASA, `http://lheawww.gsfc.nasa.gov/users/audley/diss/node31.html`.

[Bayer, 1976] Bayer, B. E. (1976). *Color Imaging Array*. United States Patent 3,971,065.

[Cok, 1987] Cok, D. R. (1987). *Signal Processing method and apparatus for producing interpolated chrominance values in a sampled color image signal*. United States Patent 4,642,678.

[Company, 2004] Company, E. K. (2004). Image sensor solutions. Technical report, Eastman Kodak Company, `http://www.kodak.com/global/en/digital/ccd/products/cmos/KAC-0311/indexKAC-0311.jhtml?id=0.1.10.6.5&lc=en`.

[Cutin, 2004] Cutin, D. P. (2004). Image sensors - capturing the photograph. Technical report, ShortCourses.Com, `http://www.shortcourses.com/how/sensors/sensors.htm`.

[Dudgeon and Mersereau, 1984] Dudgeon, D. E. and Mersereau, R. M. (1984). *Multidimensional Digital Signal Processing*. Prentice-Hall Signal Processing Series. Prentice-Hall, 1st edition.

[FillFactory, 2005] FillFactory (2005). Color filter arrays. Technical report, Fill Factory Image Sensors, `http://www.fillfactory.com/htm/technology/htm/rgbfaq.htm`.

[Freeman, 1988] Freeman, W. T. (1988). *Median Filter for reconstructing missing color samples*. United States Patent 4,724,395.

[Gonzalez and Woods, 2003] Gonzalez, R. C. and Woods, R. E. (2003). *Digital Image Processing*. Pearson Education, 2nd edition.

[Hamilton and Adams, 1997] Hamilton, J. F. and Adams, J. E. (1997). *Adaptive color plan interpolation in single sensor color electronic camera*. United States Patent 5,629,734.

[Laroche and Prescott, 1994] Laroche, C. A. and Prescott, M. A. (1994). *Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients*. United States Patent 5,373,322.

[Longere et al., 2002] Longere, P., Zhang, X., Delahunt, P. B., and Brainard, D. H. (2002). Perceptual assessment of demosaicing algorithm performance. In *Proceedings of IEEE*, volume 90, pages 123–132.

[Miao et al., 2003] Miao, L., Qi, H., and Snyder, W. (2003). A generic method for generating multi-spectral filter arrays. *IEEE International Conference on Image Processing (ICIP)*.

[Mullen, 1985] Mullen, K. (1985). The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *Journal of Physiology*, 359(1):381–400.

[Qi, 1999] Qi, H. (1999). *A High-Resolution, Large-Area, Digital Imaging System*. PhD thesis, North Carolina State University.

[Ramanath et al., 2002] Ramanath, R., Snyder, W. E., Bilbro, G. L., and Sander III, W. A. (2002). Demosaicking methods for bayer color arrays. *Journal of Electronic Imaging*, 11(3):306–315.

[Ramanath et al., 2004] Ramanath, R., Snyder, W. E., and Qi, H. (2004). Mosaic multispectral focal plane array cameras. *Infrared Technology and Applications XXX at SPIE Defense and Security Symposium.*

[Roberts, 2003] Roberts, M. (2003). *Signals and Systems: Analysis Using Transform Methods and Matlab.* Tata McGraw-Hill Publishing Company Limited, 1st edition.

[Rosenfeld and Kak, 1982] Rosenfeld, A. and Kak, A. C. (1982). *Digital Picture Processing*, volume 1. Academic Press Inc., 2nd edition.

[Rosenfled, 1970] Rosenfled, A. (1970). Connectivity in digital pictures. *Journal of the ACM*, 17.

[Snyder, 1991] Snyder, W. (1991). *Image File System Reference Manual.* NC State University Image Analysis Laboratory Database, 5.1 edition.

[Snyder et al., 2000] Snyder, W. E., Qi, H., Head, J., and Wang, C. X. (2000). Increasing the effective resolution of thermal infrared images. *IEEE Engineering in Medicine and Biology Magazine*, 19(3):63–70.

[Sonka, 2004] Sonka, M. (2004). Class notes on image pre-processing: Image restoration. Technical report, University of Iowa, `http://www.icaen.uiowa.edu/~dip/LECTURE/PreProcessing4.html`.

[Wensel et al., 1990] Wensel, D., Eberly, D., and Longbotham, H. (1990). Hexagonal tessellations in image algebra. *Image Algebra and Morphological Image Processing, SPIE*, 1350:25–30.

[Zhu et al., 1999] Zhu, W., Parker, K., and Kriss, M. A. (1999). Color filter arrays based on mutually exclusive blue noise patterns. *Journal of Vision Communication and Image Representation*, 10(3):245–267.

Vita

Gaurav Ashok Baone was born in Nagpur, India, on March 12, 1982. He completed his elementary education in Hyderabad, India. Pursuing his keen interest in mathematics, he went on to do his engineering. He graduated from Osmania University, India in 2003 with a Bachelor of Engineering degree in Electronics and Communications Engineering. During his undergraduate study, he implemented a new Anti-Slip and Anti-Slide system for Diesel-Electric Locomotives. This work was done at Bharat Heavy Electrical Limited, R&D, Hyderabad, India. In Fall 2003, he came to The University of Tennessee as a graduate student in Electrical and Computer Engineering department. He then joined the Advanced Imaging and Collaborative Information Processing lab of Dr. Hairong Qi with a major topic in Multi-Spectral Image Processing. He explored new methods for implementing the mosaic focal plane array technology for multi-spectral image acquisition systems. He developed demosaicking algorithms for multi-spectral images. He developed a novel image restoration based approach to perform demosaicking in the presence of noise and degradations. He was also involved in the development of mobile robots that were provided intelligence to move around a closed enclosure without human intervention. The robot development was a part of a successful initial venture of AICIP lab into the field of computer vision. He took a course in microelectronic systems design which added an additional aspect to his interests. He was a part of a team which designed a FPGA based handwritten digit recognition processor. His major interests and research areas are image processing, computer vision, embedded systems and digital systems design.