Mississippi State University
Scholars Junction

Theses and Dissertations | Theses and Dissertations

8-1-2005

Multiresolution based, multisensor, multispectral image fusion

Pushkar S. Pradhan

Follow this and additional works at: https://scholarsjunction.msstate.edu/td

Recommended Citation
Pradhan, Pushkar S., "Multiresolution based, multisensor, multispectral image fusion" (2005). *Theses and Dissertations*. 3237.
https://scholarsjunction.msstate.edu/td/3237

This Dissertation - Open Access is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.

www.manaraa.com

MULTIRESOLUTION BASED, MULTISENSOR, MULTISPECTRAL IMAGE FUSION

By

Pushkar Pradhan

A Dissertation
Submitted to the Faculty of
Mississippi State University
In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Computer Engineering
in the Department of Electrical and Computer Engineering

Mississippi State, Mississippi

August 2005

www.manaraa.com

Copyright by

Pushkar Pradhan

2005

www.manaraa.com

MULTIRESOLUTION BASED, MULTISENSOR, MULTISPECTRAL IMAGE FUSION

By

Pushkar Pradhan

Approved:

Roger L. King William Giles Distinguished Professor Electrical and Computer Engineering (Major Advisor)	Ioana Banicescu Associate Professor Computer Science and Engineering (Committee Member)
James Fowler Associate Professor Electrical and Computer Engineering (Committee Member)	Yul Chu Assistant Professor Electrical and Computer Engineering (Committee Member)
Nicolas Younan Professor Electrical and Computer Engineering (Committee Member and Graduate Coordinator)	Kirk Schulz Dean of Bagley College of Engineering

www.manaraa.com

Name: Pushkar Shrikant Pradhan

Date of Degree: August 6, 2005

Institution: Mississippi State University

Major Field: Computer Engineering

Major Professor: Dr. Roger L. King

Title of Study: Multiresolution based, Multisensor, Multispectral Image Fusion

Pages in Study: 192

Candidate for the Degree of Doctor in Philosophy

Spaceborne sensors, which collect imagery of the Earth in various spectral bands, are limited by the data transmission rates. As a result the multispectral bands are transmitted at a lower resolution and only the panchromatic band is transmitted at its full resolution. The information contained in the multispectral bands is an invaluable tool for land use mapping, urban feature extraction, etc. However, the limited spatial resolution reduces the appeal and value of this information.

Pan sharpening techniques enhance the spatial resolution of the multispectral imagery by extracting the high spatial resolution of the panchromatic band and adding it to the multispectral images. There are many different pan sharpening methods available like the ones based on the Intensity-Hue-Saturation and the Principal Components Analysis transformation. But these methods cause heavy spectral distortion of the multispectral images. This is a drawback if the pan sharpened images are to be used for classification based applications.

www.manaraa.com

In recent years, multiresolution based techniques have received a lot of attention since they preserve the spectral fidelity in the pan sharpened images. Many variations of the multiresolution based techniques exist. They differ based on the transform used to extract the high spatial resolution information from the images and the rules used to synthesize the pan sharpened image. The superiority of many of the techniques has been demonstrated by comparing them with fairly simple techniques like the Intensity-Hue-Saturation or the Principal Components Analysis. Therefore there is much uncertainty in the pan sharpening community as to which technique is the best at preserving the spectral fidelity. This research investigates these variations in order to find an answer to this question.

An important parameter of the multiresolution based methods is the number of decomposition levels to be applied. It is found that the number of decomposition levels affects both the spatial and spectral quality of the pan sharpened images. The minimum number of decomposition levels required to fuse the multispectral and panchromatic images was determined in this study for image pairs with different resolution ratios and recommendations are made accordingly.

www.manaraa.com

DEDICATION

I would like to dedicate this dissertation to my parents Alka and Shrikant.

www.manaraa.com

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor Dr. Roger L. King for his guidance and support throughout this research. I would also like to thank my committee members Dr. Ioana Banicescu, Dr. Yul Chu, Dr. James Fowler, and Dr. Nicolas Younan for serving on my committee and providing valuable feedback and advice.

During this research I also collaborated with different researchers working in the field of image fusion – Dr. Derrold Holcomb, Leica Geosystems, Inc., Atlanta, Dr. Ute Gangkofner, Geoville, Austria, and George Lemeshewsky, USGS. I would like to extend my sincerest thanks to them for sharing their knowledge with me. I would also like to thank my colleague Veera for the productive discussions that I have had with him regarding wavelets and pan sharpening.

Finally I would also like to acknowledge my deep gratitude to NASA and USDOT for providing financial support throughout the research.

www.manaraa.com

TABLE OF CONTENTS

www.manaraa.com

CHAPTER Page

www.manaraa.com

CHAPTER Page

www.manaraa.com

CHAPTER Page

www.manaraa.com

www.manaraa.com

LIST OF TABLES

www.manaraa.com

TABLE Page

www.manaraa.com

LIST OF FIGURES

www.manaraa.com

www.manaraa.com

CHAPTER I

INTRODUCTION

1.1 Characteristics of Remote Sensing Imagery

Remote sensing images are be characterized by their spectral, spatial, radiometric, and temporal resolutions. Spectral resolution refers to the bandwidth and the sampling rate over which the sensor gathers information about the scene. High spectral resolution is characterized by a narrow bandwidth (e.g., 10 nm). Spatial resolution refers to the smallest features in the scene that can be separated (resolved). The radiometric resolution refers to the dynamic range or the total number of discrete signals of particular strengths that the sensor can record. A larger dynamic range for a sensor results in more details being discernible in the image. The Landsat 7 sensor records 8-bit images; thus it can measure 256 unique gray values of the reflected energy while Ikonos-2 has an 11-bit radiometric resolution (2048 gray values). In other words, an higher radiometric resolution allows for simultaneous observation of high and low contrast objects in the scene. The temporal resolution refers to the time elapsed between consecutive images of the same ground location taken by the sensor. Satellite based sensors, based on their orbit, may dwell continuously on an area or revisit the same area every few days. The temporal characteristic is helpful in monitoring land use changes [22].

www.manaraa.com

Due to system tradeoffs related to data volume and signal to noise ratio (SNR) limitations, remote sensing images tend to have either a high spatial resolution and low spectral resolution or vice versa [31]. The following section explains the relationship between the spatial resolution and spectral resolution.

1.2 Spatial Resolution and Spectral Resolution Tradeoffs

All sensors have a fixed signal to noise ratio that is a function of the hardware design. The energy reflected by the target must have a signal level large enough for the target to be detected by the sensor. The signal level of the reflected energy increases if the signal is collected over a larger instantaneous field of view (IFOV) or if it is collected over a broader spectral bandwidth. Collecting energy over a larger IFOV reduces the spatial resolution while collecting it over a larger bandwidth reduces its spectral resolution. Thus, there is a tradeoff between the spatial and spectral resolutions of the sensor. As noted above, a high spatial resolution can accurately discern small or narrow features like roads, automobiles, etc. A high spectral resolution allows the detection of minor spectral changes, like those due to vegetation stress or molecular absorption [31].

Most optical remote sensing satellites carry two types of sensors – the panchromatic and the multispectral sensors. The multispectral sensor records signals in narrow bands over a wide IFOV while the panchromatic sensor records signals over a narrower IFOV and over a broad range of the spectrum. Thus, the multispectral (MS) bands have a higher spectral resolution, but a lower spatial resolution compared to the associated panchromatic (PAN) band, which has a higher spatial resolution and a lower spectral resolution. Hyperspectral sensors have been launched in recent years, which have

www.manaraa.com

even narrower channel bandwidths than MS sensors (approximately 10 nanometers), and correspondingly their spatial resolution is poorer than MS sensors (assuming a similar orbit and optics). Table 1-1 lists the spectral and spatial resolutions of the Landsat 7 ETM+ sensor [17], Table 1-2 gives the same information for the SPOT5 sensor [41]. The Landsat 7 MS bands 1-5 have a spatial resolution of 30 m while the PAN band has a 15 m resolution. Similarly, the MS bands 1-4 of SPOT5 have a spatial resolution (10 m), which is four times that of the PAN band (2.5 m). Other examples of satellite systems which carry high spatial resolution and high spectral resolution sensors are Quickbird, Ikonos, and the Indian Remote Sensing Satellite (IRS).

Table 1-1 Spectral, and spatial resolutions of Landsat 7

Band Number	Spectral Range (nm)	Spatial Resolution (m)
1	450-515	30
2	525-605	30
3	630-690	30
4	750-900	30
5	1550-1750	30
6	1040-1250	60
7	2090-2350	30
Panchromatic	520-900	15

www.manaraa.com

Table 1-2 Spectral, and spatial resolutions of SPOT 5

Band Number	Spectral Range (nm)	Spatial Resolution (m)
1	500-590	10
2	610-680	10
3	780-890	10
4	1580-1750	20
Panchromatic	480-710	2.5

1.3 Image Fusion

Image fusion refers to the process of combining two or more images into one composite image, which integrates the information contained within the individual images [5]. The result is an image that has a higher information content compared to any of the input images. Generally, the goal of the fusion process is to evaluate the information at each pixel location in the input images and retain the information from that image which best represents the true scene content or enhances the utility of the fused image for a particular application.

Researchers and other customers who buy satellite imagery desire both high spatial and spectral resolutions simultaneously in order to extract the maximum information content from the imagery. Thus, the information from both the PAN and MS channels needs to be integrated into one channel. Many different image fusion methods

www.manaraa.com

are found in the literature to combine MS and PAN images [1], [9]. These techniques combine the spatial details from a high spatial resolution-low spectral resolution (PAN) image with the low spatial resolution-high spectral resolution (MS) image to create a high spatial, high spectral resolution image. If successful, this method provides spectrally and spatially higher information content [37]. In the remote sensing literature**,** this image fusion process is popularly referred to as "pan sharpening" since the details of the PAN image are used to "sharpen" the MS imagery. A more specific and accurate definition of the pan sharpening would be "the enhancement of the spatial resolution of a low spatial resolution image by the integration of higher resolution details from an available higher spatial resolution image". This general terminology will be used since the higher spatial resolution image does not necessarily have to be that of a panchromatic band. Thus, a Landsat 7 MS image of 30 m spatial resolution could be fused with any one MS band of SPOT 5 10 m spatial resolution to increase the Landsat 7 MS image's resolution by a factor of three. One could also use a higher resolution aerial photograph of the area if it were available.

Image fusion a vast discipline in itself, refers to the fusion of various types of imagery that provide complementary information. For example, thermal and visible images are combined to aid in aircraft landing [35]. Multispectral images are combined with radar imagery because of the ability of the radar to "see" through cloud cover. Since pan sharpening refers to the fusion of only a specific type of imagery, it can be considered a subset of the vast image fusion discipline. In this study, we are only interested in applying it to the fusion of PAN and MS images. However, before image

www.manaraa.com

fusion or pan sharpening can be applied to a set of images they must satisfy certain conditions [24] –

- The images should be of the same ground location, and they must be coregistered to each other within sub pixel accuracy,
- No major changes should have occurred in the landscape imaged between the acquisition of the two images,
- The spectral ranges of the PAN and the MS images must overlap.

1.4 Applications of Pan Sharpening

Many applications such as mapping of land use, vegetation and urban areas benefit from pan sharpening. The different objects or classes observed in the scene can be better distinguished or classified due to the high spectral resolution of multispectral images. However, the maps created will have a coarse appearance due to the low spatial resolution of the MS image. On the other hand, the different classes cannot be separated in the panchromatic imagery as they have almost identical gray values [50], but the higher spatial resolution, of the panchromatic band leads to a more accurate delineation of the structures and the boundaries between them [14], [29]. Since the pan sharpened image has both a high spectral resolution and a high spatial resolution, the objects can be classified efficiently as well as delineated with higher accuracy. The maps created from the pan sharpened images leads to enhanced visual interpretation. Zhang et al. [50] used a pan sharpened image obtained from the fusion of Quickbird MS and PAN images to extract road networks in Fredericton, New Brunswick, Canada. The same principles can be used to extract buildings in an urban environment. A few papers assessing the

www.manaraa.com

effectiveness of pan sharpening for urban mapping are listed in [10], [11], and [42].

1.5 Motivation

The previous section shows that there are various applications of pan sharpening ranging from land use mapping to road extraction. All these applications involve classification of the imagery. In order that the pan sharpened imagery is classified correctly, the spectral information or the radiometry of the MS imagery must be preserved. One target application is the supervised classification of the pan sharpened imagery by using the spectral signatures derived from the original MS imagery [12].

Thus, preserving the spectral information of the original MS images in the pan sharpened images is important. This means that there should be ideally zero or minimal change in the radiometry or digital number (DN) values of the image. The change or loss of the original radiometry is also termed as 'spectral distortion' in literature. Over the years, researchers have formulated various techniques for pan sharpening that attempt to minimize the spectral distortion, or in other words, retain the maximum spectral fidelity of the MS images [1], [7], and [14]. On the other hand, if the use of the pan sharpened image is just to produce maps for better visual interpretation, the spectral distortion is not of much concern, in that case, the goal is then to produce images with high contrast.

The goal of this dissertation is to produce pan sharpened images with the highest spectral fidelity possible because of the importance of such images in classification tasks. The literature survey and my initial experiments showed that Multi Resolution Analysis (MRA) based methods seem to be the most effective at producing high spectral fidelity pan sharpened images. Thus, the MRA techniques are pursued and investigated in detail.

www.manaraa.com

There are many variants of the MRA techniques, and each is promoted as being the best at preserving the spectral fidelity of the MS images. Arguments are made in favor of a certain new technique by comparing it against fairly simple techniques like the IHS or PCA methods. One of the novel approaches of this work is the rigorous comparison of different MRA methods with each other.

1.6 Contributions

The main goal of this study is to implement MRA based pan sharpening techniques and its many variants and determine empirically the most suitable ones. The motivation behind pursuing MRA based techniques is that they preserve the spectral fidelity of the MS images while improving their spatial resolution. A number of parameters of the MRA based pan sharpening methods were studied, and their effects on the quality of the pan sharpened images determined quantitatively. The major findings of this work are described in the following subsections.

1.6.1 Choice of the Mother Wavelet

The most suitable mother wavelet for wavelet based pan sharpening is the second order Daubechies wavelet ‘db2’ based on empirical evaluations. A more general inference is that the shorter wavelets gave better results than the longer ones. The only exception was the shortest 2-tap Haar wavelet, which gave poorer results compared to even the longest wavelets. Similarly, various classes of filters are used for the Laplacian Pyramid (LP), and it is determined that the 13-tap Quadrature Mirror Filter (QMF) gave the highest spectral fidelity. The same filters that are used for the LP based pan

www.manaraa.com

sharpening were also used to pan sharpen images with the A Trous Wavelet Transform (AWT). It is determined that the Gaussian filter corresponding to a central coefficient weight (CCW) of 0.5 gave the highest spectral fidelity among all the filters explored.

1.6.2 Comparison of Various MRA Transforms

The various MRA transforms like the Redundant Wavelet Transform (RWT), the A Trous Wavelet Transform (AWT) and the Laplacian Pyramid (LP) are compared. In the many pan sharpening papers found in literature, there are different and contradictory claims regarding which one is the best. A detailed comparison between RWT, AWT, and LP is carried out. As a result, it is determined that the RWT is best at preserving the spectral fidelity of the pan sharpened image.

1.6.3 Benefits of the Additive Method

It is shown that the additive pan sharpening technique performs well when the MRA transform is a LP and the spectral fidelity provided by this algorithm is quite close to the substitutive method based on the RWT. The advantage of this technique is that, since only the PAN image has to be decomposed, it is computationally faster compared to the substitutive method. However, sometimes the spatial quality of this method is slightly less than that of the substitutive methods.

1.6.4 Choice of the Selection Rule

There are various selection rules proposed to combine the high frequency coefficients of the MS-PAN images in order to improve the spectral fidelity of the pan

www.manaraa.com

sharpened images. These rules are evaluated and it is found that the maximum amplitude selection rule provides the highest spectral fidelity while being computationally the least expensive.

1.6.5 Effect of Directional Selectivity

Often the reason for the superiority of the RWT over the LP or the AWT is cited as the higher directional selectivity of the RWT. Through experiments based on the Steerable Pyramid Transform (SPT), whose directional selectivity can be varied, it is determined that the directional selectivity played only a minor role in improving the spectral quality.

1.6.6 Estimation of the Minimum Number of Decomposition Levels

The relationship between the number of decomposition levels required to fuse PAN and the MS images and the resolution ratio of the images is found. The number of decomposition levels required to merge images is directly proportional to their resolution ratio. Recommendations on the minimum required decomposition levels for a given resolution ratio are made, based on the pan sharpening application. It is found that the spatial quality is affected by the number of decomposition levels, and other parameters of the fusion process have hardly any effect on it, i.e. a different wavelet filter or selection rule cannot reduce the number of decomposition levels required to merge the datasets.

www.manaraa.com

1.7 Document Overview

The next chapter (Chapter 2) describes the various pan sharpening methods found in literature with an emphasis on MRA based techniques. The reader is also introduced to the theory of multiresolution analysis. Chapter 3 describes the metrics used for quantitative assessment of the pan sharpened images. The topics that will be researched are also outlined along with the methodologies used to approach them. Chapter 4 describes the datasets that were used in this research, while Chapter 5 gives detailed results on each of the topics that were investigated. Chapter 6 makes conclusions based on the results obtained and suggestions for future work. Appendix A defines some acronyms used throughout the document while in Appendix B some of the MRA transforms are explained.

www.manaraa.com

CHAPTER II

LITERATURE REVIEW

2.1 Introduction

There are many pan sharpening techniques available in literature. This chapter describes some popular ones; any advantages or disadvantages associated with them are also given. Review articles on image fusion have appeared in literature describing the various techniques available [28]. The authors of [28] classify pan sharpening techniques into categories such as arithmetic combination based, color based, wavelet based, and substitution based methods since many of them are minor variations of a pioneering technique. The wavelet based methods are just one of the many realizations possible of the MRA based pan sharpening technique. In this work I prefer to use a slightly different categorization scheme, the one proposed by Shettigara [36] in which, any technique that involves a forward transformation of the bands, replacement or manipulation of the transformed bands and a reverse transformation is called a Component Substitution (COS) technique. Figure 2-1 shows the categorization of pan sharpening techniques that I follow and a few techniques that fall under them. Most pan sharpening methods can be classified as COS based or arithmetic combination based (AMC), i.e. they directly perform some type of arithmetic operation on the MS and PAN bands like weighted

www.manaraa.com

addition, multiplication, normalized division, ratios etc. According to Shettigara's approach, the MRA based techniques would be classified under the COS technique as they involve a forward and an inverse MRA transform to construct the pan sharpened images. However, the MRA based methods have received so much attention and so many variations exist that they can be considered a separate category in them self.

The following sections describe the COS, AMC and MRA methods and provide an example of each. The focus of this work is the MRA based pan sharpening methods, of which many variations exist. Thus, this work provides a comprehensive review of the MRA based methods. To explain the algorithms, throughout this report the algorithms the following notations will be used: B represents the set of MS bands, A – the PAN image and F – the pan sharpened bands. The notations B and F will be used to represent all the bands collectively. If a particular band is to be addressed a subscript shall be used, e.g. B_1 refers to the first band in the MS set. Most of the techniques require that the MS images first be resampled to the pixel size of the PAN image, thus the resampled MS images will be noted by the superscript *, i.e. B^*. Any image downsampled or degraded to a lower resolution from its original resolution will be denoted by the superscript `, i.e. A` refers to the PAN image degraded to a lower resolution.

www.manaraa.com

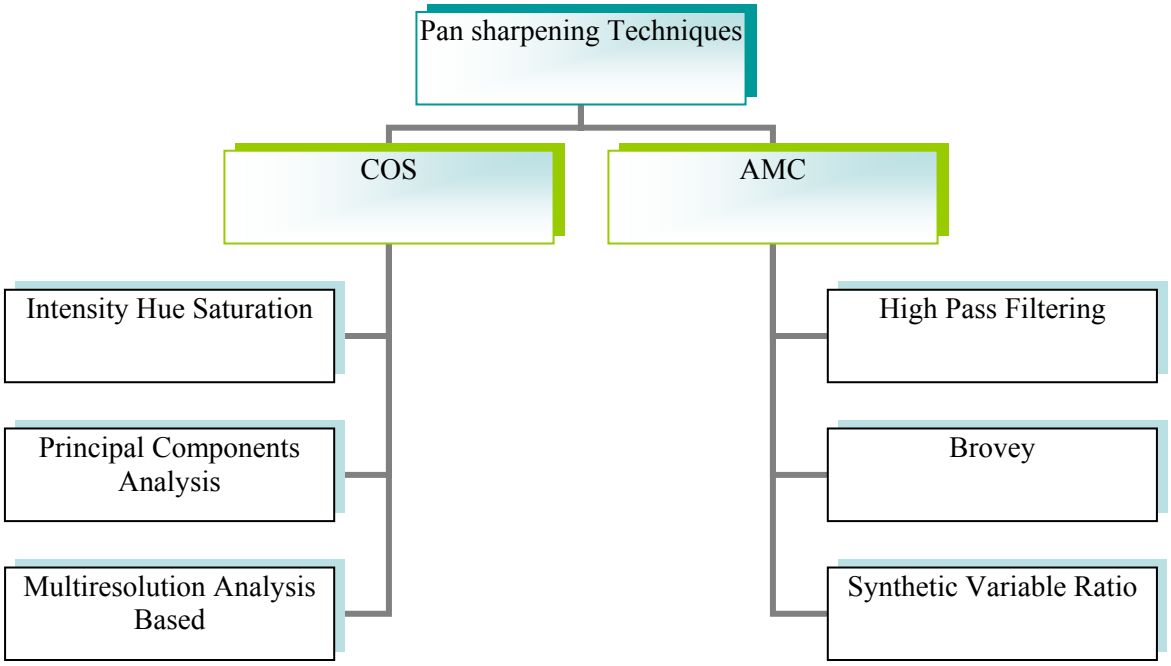


Figure 2-1 Categorization of pan sharpening techniques

2.2 The COS Methods

Some of the popular COS methods for pan sharpening are the IHS method, the PCA method, and the MRA based methods. Grasping the working of MRA based methods, requires a proper understanding of the MRA concepts, which are quite complex. Thus, this topic is introduced at a later stage in the chapter (section 2.4), with the hope that after reading the other COS and AMC methods the reader will be more comfortable with the pan sharpening solutions and able to comprehend the MRA technique better.

www.manaraa.com

2.2.1 The Intensity Hue Saturation Method

2.2.1.1 The Intensity Hue Saturation Transform

Digital images are generally displayed by an additive color composite using the three primary colors – red (R), green (G), and blue (B) [22]. However, colors can be described by an alternate representation: Intensity, Hue, and Saturation (IHS). Intensity represents the total luminance of the image, hue represents the dominant wavelength contributing to the color, and saturation describes the purity of the color relative to gray. The IHS transformation separates the spatial and spectral information in a RGB image. The intensity component represents the spatial information while the hue and saturation describe the spectral information [22]. The spatial information can then be manipulated by performing some mathematical operation on the "intensity" component to enhance the image without altering its spectral representation. This principle is used in the IHS pan sharpening scheme described below.

More than one algorithm exists to compute the RGB to IHS transform, and they differ mainly in the way "intensity" is defined. Two of the most popular transform pairs to convert from RGB to IHS representation and back from IHS to RGB representation are the hexcone model and the triangle model. In order to distinguish between the two representations, the intensity component of the hexcone model will be called "value" (V) and that of the triangle model called "luminance" (L). Correspondingly, the IHS representation of the hexcone model will be denoted by VHS and that of the triangle model denoted by LHS. The hexcone model is described in detail below.

www.manaraa.com

2.2.1.2 The Hexcone Model

This transform pair defines intensity as the maximum value of the R, G, and B values at any pixel position. The details to calculate VHS from RGB and the reverse RGB to VHS are given below [39]:

RGB TO HSV

1) V := max(R,G,B),
2) Let X := min(R,G,B),
3) S := (V-X)/V, if S==0 return,
4) Let r := (V-R)/(V-X), g := (V-G)/(V-X), b := (V-B)/(V-X),
5) If R=V then H := (if G=X then 5+b else 1-g),

   If G=V then H := (if B=X then 1+r else 3-b),

   else H := (if R=X then 3+g else 5-r),
6) H := H/6

HSV TO RGB

1) H := H*6,
2) Let I := floor(H), F := H-I,
3) Let M := V*(1-S),

   N : = V*(1-(S*F)),

   K := V*(1-S*(1-F))),
4) Switch(I):

   Case 0: (R,G,B) := (V,K,M),

   Case 1: (R,G,B) := (N,V,M),

www.manaraa.com

Case 2: (R,G,B) := (M,V,K),

Case 3: (R,G,B) := (M,N,V),

Case 4: (R,G,B) := (K,M,V),

Case 5: (R,G,B) := (V,M,N),

In the above transforms the RGB space is first normalized to the range [0, 1] and thus all RGB and HSV values fall in this range. A point to note is that if the saturation (S) is zero it means all R, G, B are equal at this point, and the hue is not defined at this point. The hue is then replaced by the immediate previous value in the image.

2.2.1.3 The Triangle Model

The other transform pair is called the triangle model, which defines the intensity as the average of R, G, and B at the pixel position: L = (R+G+B)/3. The interested reader can find the details of this transform pair in [39].

2.2.1.4 IHS Based Pan Sharpening

The IHS pan sharpening technique is the oldest known data fusion method and one of the simplest. Figure 2-2 illustrates this technique for convenience. In describing the algorithm the nonspecific notation IHS will be used since either VHS or LHS can be used for the implementation. In this technique the following steps are performed:

1. The low resolution MS imagery is co-registered to the same area as the high resolution PAN imagery and resampled to the same resolution as the PAN imagery.

www.manaraa.com

2. The three resampled bands of the MS imagery B^*, which represent the RGB space are transformed into IHS components.
3. The PAN imagery is histogram matched to the "I" component. This is done in order to compensate for the spectral differences between the two images, which occurred due to different sensors or different acquisition dates and angles.
4. The intensity component of MS imagery is replaced by the histogram matched PAN imagery. The RGB of the new merged MS imagery is obtained by computing a reverse IHS to RGB transform.

In the above algorithm, replacing the spatial component of the MS imagery with the PAN imagery allows the details of the PAN imagery to be incorporated in to the MS imagery.

www.manaraa.com

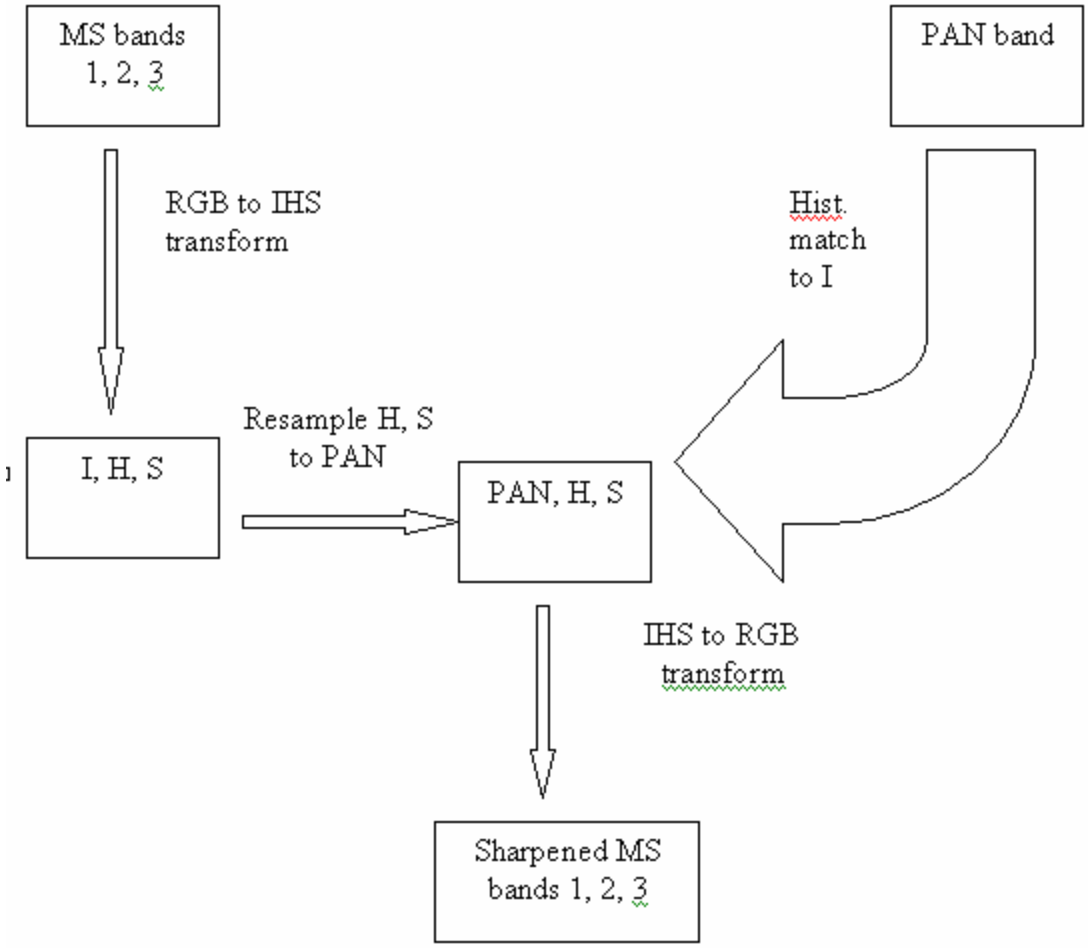


Figure 2-2 IHS based pan sharpening method

2.2.1.5 Comments on the IHS Method

The IHS technique is fairly easy to understand and implement. Moreover it requires very little computation time compared to the more sophisticated techniques. However, it severely distorts the spectral values of the original color of the MS image. This is seen from Figure 2-3 – Figure 2-5, which show a QuickBird image pan sharpened using the VHS method. The colors of the buildings in the lower right region of the image

www.manaraa.com

(below the diagonal road) have changed from white or light blue to a strong blue. The vegetation above the road also appears lighter in color. Thus the IHS technique is good only for visual analysis, not machine classification; moreover it is also limited to three bands.

Figure 2-3 Quickbird MS image

www.manaraa.com

Figure 2-4 Quickbird PAN image

www.manaraa.com

Figure 2-5 The VHS pan sharpened image

Since there are different techniques to compute the IHS representation, it is important to make comparisons among them. A brief literature survey of IHS based pan sharpening papers showed that [7], [26] used the LHS representation. Nunez et al. [26] gave quantitative results showing that the LHS method resulted in less spectral distortion of the MS bands compared to VHS.

www.manaraa.com

2.2.2 Principal Components Analysis Method

2.2.2.1 Principal Components Analysis – A Dimensionality Reduction Technique

The data in each of the MS bands of a sensor are highly correlated. Thus, the different bands are quite similar and the data is redundant to some extent. If an analyst must process this data, the redundant information must be filtered for effective analysis. The goal of the principal components analysis is to reduce the dimensionality of the data by decorrelating the original data in N dimensions into uncorrelated data in fewer than N dimensions [22]. This is analogous to data compression.

The principal components (PC) of an N dimensional data are computed by a linear transformation [8]. Suppose x is an Nx1 vector, then the mean of x can be estimated from a sample of L such vectors:

$$m_x = \frac{1}{L}\sum_{l=1}^{L} x_l \quad , \tag{1}$$

and its covariance matrix is given by:

$$C_x = E\{(x-m_x)(x-m_x)^T\} = \frac{1}{L}\sum_{l=1}^{L} x_l x_l^T - m_x m_x^T \quad , \tag{2}$$

where E is the expectation operator. The covariance matrix is a NxN matrix and it is real and symmetric. Its diagonal elements are the variances of the individual variables and the off-diagonal elements are their covariances. Then a matrix A is defined such that its rows are eigenvectors of C_x, and it defines a new vector y, which is a linear transformation of x:

www.manaraa.com

$$y = A(x - m_x), \tag{3}$$

The new vector y is a vector with zero mean and its covariance C_y is related to C_x as follows:

$$C_y = AC_xA^T, \tag{4}$$

Since the rows of A are eigenvectors of C_x, C_y is a diagonal matrix with the eigenvalues of C_x as its elements:

$$C_y = \begin{bmatrix} \lambda_1 & . & . & 0 \\ 0 & \lambda_2 & . & 0 \\ . & . & . & . \\ 0 & . & . & \lambda_N \end{bmatrix}, \tag{5}$$

Since the off-diagonal elements of C_y are zero, the variables of the transformed vector y are uncorrelated and independent of each other. The rows of A are arranged in order of decreasing magnitude, and thus the variables of y are arranged in order of decreasing significance. For dimensionality reduction we can neglect one or more elements of y that correspond to the smallest magnitude eigenvectors. This transform is invertible; this operation can be called the inverse principal components:

$$x = A^{-1}y + m, \tag{6}$$

2.2.2.2 PCA Pan Sharpening

In the PCA pan sharpening method, instead of calculating the IHS transform of the MS imagery, the principal components are calculated. In this technique the following steps are performed, Figure 2-6 also illustrates these steps:

www.manaraa.com

1. The low resolution MS imagery is co-registered to the same area as the high resolution PAN imagery.
2. The principal components of the MS imagery are calculated and resampled to the same pixel size as the PAN image.
3. The PAN image is histogram matched to the first principal component (PC1) of the MS image. This matching is done in order to compensate for the albedo differences between the two images.
4. The first PC of the MS imagery is replaced by the histogram matched PAN imagery, which has a higher resolution. The logic behind replacing the first PC is that it contains the information common to all the bands. So replacing it with the high resolution PAN image will maximize the effect of the fusion in all the bands [36].
5. The inverse principal components are computed to obtain the new merged MS imagery.

www.manaraa.com

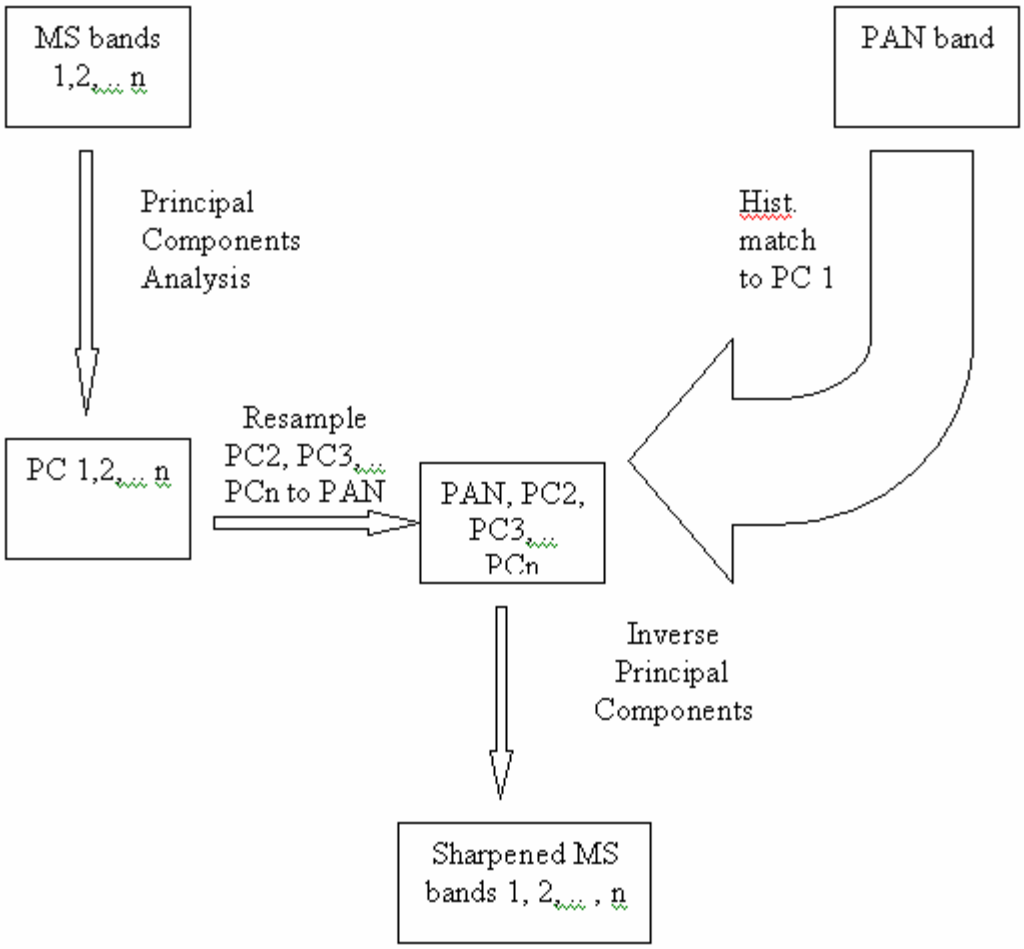


Figure 2-6 PCA based pan sharpening method

2.2.2.3 Comments on the PCA Method

An advantage of this technique is that the number of input bands is theoretically unlimited. The PCA method for pan sharpening gives better results compared to the IHS method. However, this method is sensitive to the area being sharpened [28] because the variance between pixel values in a band and the correlation among the MS bands differs based on the land cover. Moreover, since the first PC contains the maximum variance of the MS imagery, replacing it with the PAN band maximizes the spectral effect of the

www.manaraa.com

PAN image. Figure 2-7 shows the above QuickBird MS image pan sharpened using the PCA method. The spectral quality of the image is not preserved. The buildings in the lower left region below the road appear bluish instead of white as in the original MS image. The bare soil and grass on the right side of the road appear brighter and light green respectively.

Figure 2-7 PCA based pan sharpened image

www.manaraa.com

2.3 The AMC Methods

The logic behind the AMC methods is to increase the spatial resolution of the MS imagery by adding, multiplying or taking some sort of ratio of the PAN image with them [28]. To minimize the spectral effects of the PAN image most implementations do a weighted addition or multiplication, where the weights are determined based on some criteria.

2.3.1 Addition, Multiplication

The following two formulas were proposed by Yesou et al. [45] to enhance the spatial resolution of the MS images,

$$F_i = s_A(w_1 A + w_2 B_i^*) + s_B, \tag{7}$$

$$F_i = s_A(AB_i^*) + s_B, \tag{8}$$

The additions and multiplications are point wise operations. s_A, s_B are the scaling factors and w_1, w_2 are the weighting factors. The MS images have to be first resampled before applying the above formulas. Many other arithmetic possibilities exist and the references for these can be found in [28]. But, since these methods do not preserve the spectral fidelity of the MS images well, they are not considered in this work.

2.3.2 Ratioing

Certain band ratios enhance the image details and thus, this technique is very suitable for producing images with high contrast. The Brovey transform and the Synthetic Variable Ratio (SVR) methods are examples of this technique.

www.manaraa.com

2.3.2.1 The Brovey Transform

The Brovey transform method (named after its author) uses ratios to sharpen the images. Each MS band is normalized with respect to the other MS bands and multiplied by the PAN band to obtain the pan sharpened image. The following equation gives the mathematical formula for the Brovey method [12]:

$$F_i = \frac{B_i^*}{\sum_{j=1}^{N} B_j^* + c} \cdot A, \tag{9}$$

In practice there can be areas in the image where the DN values in all the bands are zero, thus a small constant c is added to the denominator to produce meaningful output values. This method produces images with excellent contrast, which can be useful for change detection. However, the pixel values are altered drastically and thus these images are not suitable for automatic classification purposes. Figure 2-8 shows that MS image pan sharpened by the Brovey method has a high contrast. The bare region in the top of the image has changed from brown to blue. The road, which goes across the image diagonally, has also changed from grayish-blue to a brown color. Moreover, a comparison with the pan sharpened images from the IHS and PCA methods shows that the Brovey images introduce the most spectral distortion.

www.manaraa.com

Figure 2-8 Pan sharpening by Brovey method

2.3.2.2 The Synthetic Variable Ratio Method

The original Synthetic Variable Ratio (SVR) method was proposed by Munechika et al. [25], which is given by the following equation:

www.manaraa.com

$$F_i = \frac{B_i}{A'_{syn}} \cdot A, \quad (10)$$

where A'_{syn} is the low resolution PAN image created synthetically from the low resolution MS bands. The aim is to create a low resolution PAN image that would have been measured by a low resolution panchromatic sensor if it was available. This image is calculated as follows:

$$A^i_{syn} = \sum_{i=1}^{N} \phi_i B_i \ , \quad (11)$$

The coefficients ϕ_i are calculated by using an atmospheric model. The representative reflection spectra of five land cover classes – urban, soil, water, trees and grass in the image under consideration are measured. This spectra and the target reflectance are input to an atmospheric model that simulates the reflectance values for the PAN band. The ϕ_i coefficients are then obtained through regression analysis between the simulated and actual DN values. Finally, before applying equation $F_i = \frac{B_i}{A'_{syn}} \cdot A, \quad (10)$ the PAN image is histogram matched to the synthetic low resolution PAN image. Zhang [48] notes that the coefficient calculation is complicated as it requires measuring the spectra of particular classes and running an atmospheric model that requires considerable human operator skills. Moreover, many scenes may not have these five specified classes, there may be more than these classes. Thus, this situation creates additional work for the operator.

www.manaraa.com

Zhang proposed a modified SVR method that calculates the regression analysis coefficients ϕ_i directly from the original PAN (A) and the resampled MS images (B^*). These coefficients are then used to calculate the synthetic PAN image – A_{syn} at the same resolution as the original PAN image given by:

$$A_{syn} = \sum_{i=1}^{N} \phi_i B_i^* , \tag{12}$$

Thus the modified SVR formula is:

$$F_i = \frac{B_i^*}{A_{syn}} \cdot A , \tag{13}$$

The primary advantage of this method is that the calculation of regression analysis coefficients ϕ_i is much simpler and does not require operator intervention. Though the Zhang designed the modified SVR to preserve the spectral information of the MS images, my initial results with this method were not very encouraging (loss of spectral fidelity). A literature review showed that Zhang pursued another technique based on least squares error citing loss of spectral fidelity in the SVR method as its drawback [49].

2.3.3 The High Pass Filtering Method

The High Pass Filtering (HPF) technique is one type of a weighted addition of the PAN image with the MS image. However, because it played an important role in the realization of the MRA based methods it is discussed separately. The problem of pan sharpening can be considered as the substitution of the high spatial frequency information from one image to another [1]. The PAN image contains the high spatial frequency

www.manaraa.com

information that must be inserted or added into the MS image. The first such technique to take advantage of high spatial frequency substitution from the PAN image to the MS image was the High-Pass Filtering (HPF) method developed by Schowengerdt [34]. The HPF method is implemented as follows:

1. The low resolution MS imagery is co-registered to the same area as the high resolution PAN imagery and resampled to the same size as the PAN image using a bicubic resampling technique.
2. The high frequency details of the PAN image are obtained by subtracting a lowpass version of the PAN image from itself. The lowpass filtering is implemented by simply averaging pixels over a $n \times n$ (typical values for n are 3, 5, or 7) sliding window which moves over each pixel in the image. For example the lowpass filter used by Schowengerdt was:

$$LP = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}, \tag{14}$$

3. The high frequency details are then added to the resampled MS image band by band, which gives a high resolution MS image. Since the highpass information from one band is being substituted to another band, the difference in radiometry between the two bands must be accounted for first. Mathematically, the HPF method can be given by the following equation,

$$F_i = B_i^* + A_{hp}, \tag{15}$$

where A_{hp} is the high pass filtered PAN image.

www.manaraa.com

The main drawback of the HPF technique is the results are very much dependent on the window size over which the high frequency details image is computed. A good amount of trial and error is required in coming up with the optimal window size for a particular scene, i.e. the results vary with the landscape [28]. Nevertheless, this technique laid the foundation for advanced MRA based pan sharpening techniques. These techniques use the same principle as the HPF method – that only the high pass filtered information or the higher frequencies from the PAN image are needed to increase the spatial resolution of the MS images.

2.4 The MRA Based Methods

2.4.1 Introduction

The MRA based methods use the same principle as the HPF technique – extract the high frequency or details information from the PAN image for insertion into the MS image. The difference lies in the way in which this information is extracted. The MRA transforms perform a forward transformation to separate the low and high frequency information in the images, substitute the high frequency information from the PAN into the MS image and finally perform a reverse transform to construct the pan sharpened image. Most implementations perform a forward transform on the MS images, also [14], [18], [21], discard the high frequency information extracted from the MS image and substitute it with that from the PAN image before performing the reverse transformation.

www.manaraa.com

2.4.2 Multi Resolution Analysis Theory

Pattern recognition in remote sensing imagery encompasses the automatic identification of variable sized objects in the image. For example, if there is a requirement to detect all the edges in an image, these edges could be large or small depending on how rapidly the image intensity changes. The small edges can be easily identified by applying small neighborhood detection operators while the large edges can be identified by large neighborhood operators. However, it is recognized that using a large operator to detect a larger feature is computationally very inefficient [8]. Rather than scaling the operator to the scale of the object, it is more efficient to vary the scale of the image (or the object). Thus, the image analysis (pattern recognition) is more efficient if the image is analyzed at different resolutions [23].

An image at a given resolution can be divided into coarser approximations at a lower resolution. Suppose the original image has a resolution r_j and its lower resolution approximation image has resolution r_{j-1}. Then, the details missing in the lower resolution representation are given by the difference between the original image and the approximation image [23]. At coarser resolutions only the large objects are visible and the viewer gets only a rough idea of the image context. The original image can be reconstructed as successive details are added to the approximations and the finer details of the image become visible. A mathematical definition and an explanation of multiresolution are given below.

Multiresolution: A continuous function $f(t)$ can be decomposed into many subspaces, where each subspace contains a part of the whole function. Each of these

www.manaraa.com

different parts contains a projection of the function at different resolutions. This decomposition of the function onto subspaces at different scales or different resolutions can be defined as multiresolution [40].

In order to explain the multiresolution theory two subspaces V_j and W_j are used and j denotes the scale of the subspace. The projection of $f(t)$ on the subspace V_j is denoted by $f_j(t)$. As the scale $j \to \infty$ the projection $f_j(t)$ better approximates the function $f(t)$. It should also be noted that each subspace V_j is contained within the next higher subspace V_{j+1}, which can be generalized by the following equation:

$$V_0 \subset V_1 \subset \cdots V_j \subset V_{j+1} \subset \cdots, \tag{16}$$

The details missing in V_j to construct V_{j+1} are contained in the subspace W_j, thus the subspace V_{j+1} can also be written as:

$$V_{j+1} = V_j \oplus W_j, \tag{17}$$

The above equation can be generalized for any two consecutive subspaces (i.e., $V_1 = V_0 \oplus W_0$) which leads to the following summation:

$$V_{j+1} = V_0 \oplus W_0 \oplus W_1 \cdots W_j, \tag{18}$$

In addition to the above properties, the subspaces must also satisfy the dilation and translation requirements. The dilation requirement states that all the rescaled functions in V_j will be in V_{j+1}. If a scaling factor of two is chosen, then it means that if the subspace V_j contains frequencies up to f, V_{j+1} must contain the frequencies up to $2f$. This is given by the following equation:

www.manaraa.com

$$f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1}, \tag{19}$$

While the translation requirement states that if $f_j(t)$ is in V_j then a shifted version of $f_j(t)$ - $f_j(t-k)$ must also be in V_j.

$$f_j(t) \in V_j \Leftrightarrow f_j(t-k) \in V_j, \tag{20}$$

Finally, a scaling function $\phi(t)$ must be defined that generates the approximations of the function $f_j(t)$ on each subspace V_j. The translations of the $\phi(t)$ - $\phi(t-k)$ must span the whole space V_0 and be orthonormal. The scaling function $\phi_{j,k}(t)$ for each subspace V_j is generated by the dilations and translations of $\phi(t)$:

$$\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k), \tag{21}$$

The details that are present in the subspace W_j can be taken by the simple difference between two successive approximations $f_j(t)$ and $f_{j+1}(t)$ or by decomposing the function $f(t)$ on the wavelet function $-\psi(t)$. The wavelet functions for each subspace W_j are obtained by dilations and translations of this basic wavelet function:

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k), \tag{22}$$

This multiresolution analysis is not restricted to the continuous functions $f(t)$, but can be extended to discrete functions, which can be a one dimensional signal or a two dimensional image. The LP, AWT, DWT, and the RWT are all transforms that perform MRA, but using different approaches. Each of these transforms is described in detail in

www.manaraa.com

Appendix B for the interested reader. However, here only the basic properties of these transforms and the differences among them will be described.

Any MRA transform used generates two types of subband images – the approximation or low frequency image at the lowest scale requested and one or more details or high frequency images at each scale depending on the transform. For example, the RWT and the DWT generate three detail images (H, V, D) at each scale corresponding to the details of the image in three directions – horizontal, vertical, and diagonal, respectively. While the LP and the AWT generate only a single detail image at each scale without any spatial orientation. The DWT and RWT compute the details by decomposing the signal over the wavelet function while the LP and AWT obtain the details by simply taking the difference between the approximation images at successive scales. Further, the DWT and the LP are subsampled transforms (i.e. the images are decimated by a factor of two at each consecutive scale) while the RWT and the AWT are oversampled transforms because there is no decimation at consecutive scales.

2.4.3 Categorization of the MRA Pan Sharpening Methods

As noted earlier there are many different versions of the MRA pan sharpening method. These versions differ due to the choice of preprocessing options, the MRA transform applied, applying the MRA transform on only the PAN or both the PAN and MS images. The meaning and purpose of each of these choices is explained below.

www.manaraa.com

2.4.3.1 Radiometric Normalization

The first step in most pan sharpening methods like the IHS, PCA, and similar COS methods is to adjust the PAN band to the component it is going to replace (i.e. Intensity or PC1). This reduces the spectral effect of the PAN band in the pan sharpened image. If the PAN and MS images have been acquired on different dates under different solar angles, etc., the radiometry of the PAN will be very different from that of the MS images [26] and consequently the spectral effect of the PAN image will be severe. This adjustment or radiometric normalization is generally accomplished by a histogram matching of the PAN band to the component being replaced.

Generally the same logic is extended to MRA based pan sharpening, most researchers histogram match the PAN band to each MS band before decomposing them and substituting the high frequency coefficients of the PAN image in place of the MS image's coefficients [26], [37]. However, some implementations have also been found where the radiometric normalization is left out [18], [29], and [52]. In [18], [52] Landsat 7 TM and SPOT PAN images were fused, in which case the spectral differences between the PAN and MS should be significant since they were acquired on different dates. However, it is difficult to tell whether their methods could have benefited from the normalization step since they did not mention whether they experimented with the radiometric normalization step or not.

2.4.3.2 MS Transformation

The reason for the poor spectral fidelity of the IHS and the PCA methods was their adding redundant spatial information to the MS images by superimposing the PAN

www.manaraa.com

image in place of the Intensity or the first PC rather than adding only the high frequency spatial information. Some researchers found that instead of applying the MRA based pan sharpening to each MS band, if their IHS or PCA transformation was computed and the MRA based pan sharpening applied only to the Intensity or first PC, the spectral fidelity improved. Nunez et al. [26] applied the MRA based technique to the Intensity component of the MS images while [14] applied the MRA based pan sharpening to the first PC.

However, performing pan sharpening on only the Intensity or first PC can have drawbacks if the images contain regions of contrast reversal [19]. Contrast reversal occurs when the scene imaged has opposite contrast in the MS and PAN images. For example water has a strong reflectance at shorter wavelengths (visible band), while it absorbs heavily at longer wavelengths (NIR band). Suppose for a certain material the PAN band has low reflectance or, in other words, low contrast and the MS band has high contrast. Then if the pan sharpening is performed there are artifacts in this region because the high frequency coefficients of opposite polarity cancel each other out in the pan sharpened image resulting in loss of detail information. Thus contrast reversals must be detected in the images and special processing must be applied to those regions where contrast reversal occurs. Since this has to be done for each MS band the IHS and PCA based MRA methods are not suitable in such a scenario. Our literature survey showed that very few articles explicitly deal with the issue of contrast reversal [2], [5], [19], [34], this is perhaps because contrast reversal occurs under very few situations.

www.manaraa.com

2.4.3.3 The MRA Transform

The core principle of the MRA pan sharpening methods is the separation of the low and high frequencies in the images and substituting the high frequency information of the PAN image into the MS images. This step (the MRA transform) occurs after one or both of the above preprocessing steps are applied. Any wavelet transform or a pyramid transform can be used to accomplish this. The various MRA transforms that have been used in pan sharpening research are –

- Discrete Wavelet Transform (DWT) [20],
- Redundant Wavelet Transform (RWT) [1],
- A Trous Wavelet Transform (AWT) [26],
- Laplacian Pyramid (LP) [1],

The differences between these various transforms were given in section 2.4.2. The pan sharpened images differ depending on the transform used to obtain them. This topic is further addressed in section 3.2.6 and the Results chapter.

2.4.3.4 Coefficient Synthesis Method

Once the MS and PAN images have been decomposed into their detail and approximation coefficients there are different techniques to combine them and synthesize the pan sharpened images. Primarily, there are two variants of the MRA based method – additive and substitutive methods, explained in the following subsections. Out of the two, the substitutive method and its variants are more popular. Since each of them can be

www.manaraa.com

implemented using almost any MRA transform, the explanations will be given with a generic notation.

In the following discussion the detail coefficients (or images) obtained by applying the MRA transform will be denoted by d while the approximation image will be denoted by a. The detail coefficients of the MS image will be subscripted by B and thus denoted by d_B and those coming from the PAN image by d_A. Similarly, the approximation images from the MS and PAN images will also be subscripted by B and A respectively. The combined or selected coefficients, used to reconstruct the pan sharpened image are subscripted with F - d_F, a_F. The approximation or detail coefficients will be indexed by a vector $\vec{p} = (i, j, k, l)$ as suggested by [5]. i, j represent the pixel location in the transformed images, k denotes the level of the wavelet transform and l the orientation of the detail image. For transforms like the AWT or LP l is one since there is only one detail image at each level in the pyramid, while for the DWT or RDWT l can be one, two or three corresponding to the three detail images – H, V, D. Since any MRA transform can be used to implement a particular scheme, the acronym *MT* will be used to convey the meaning that a forward MRA transform is applied and *IMT* to mean that an inverse MRA transform is applied.

2.4.4 Additive Method

In the additive method described by Nunez et al. [26], the AWT was used to decompose only the PAN image and then the detail or high frequency coefficients are added to the MS imagery. They found that the additive method resulted in less spectral

www.manaraa.com

distortion of the MS images compared to the substitutive method (described in the following section) because all of the original information in the MS image is retained in the pan sharpened image.

The merging can be carried out on each MS band separately or only the Intensity component of the MS imagery, the latter preferred by the authors since it gives higher spectral fidelity. They found that adding the details to all the bands separately resulted in shifting the color of the pixel towards gray. Pan sharpening only the "intensity" component is also better from the performance point of view because we only need to compute the MRA transform on two images (I and PAN) rather than four images (R, G, B and PAN). Figure 2-9 shows the flowchart of the additive pan sharpening method. The main steps in the additive method are as follows:

1. Both the images are coregistered and the MS imagery is resampled to the same spatial resolution as the PAN imagery.
2. The RGB bands of the MS imagery are transformed into IHS components as described above in section 2.2.1.1.

$$(B_1^*, B_2^*, B_3^*) \xrightarrow{RGBTOIHS} = (I, H, S), \tag{23}$$

3. The PAN image is histogram matched to the Intensity component of the MS imagery (as described in section 2.4.3.1 this step is optional).
4. The detail coefficients of the histogram matched PAN image are obtained:

www.manaraa.com

$$\{a_A, d_A\} = MT(A), \tag{24}$$

5. The approximation image of the PAN band – a_A is discarded and the Intensity band of the MS image (I) substituted in its place. The inverse wavelet transform is computed on them to obtain the sharpened Intensity component I_f:

$$I_f = IMT(I, d_A), \tag{25}$$

6. The inverse IHS to RGB transform is computed to obtain the sharpened MS imagery:

$$(I_f, H, S \xrightarrow{IHSTORGB} = (F_1, F_2, F_3), \tag{26}$$

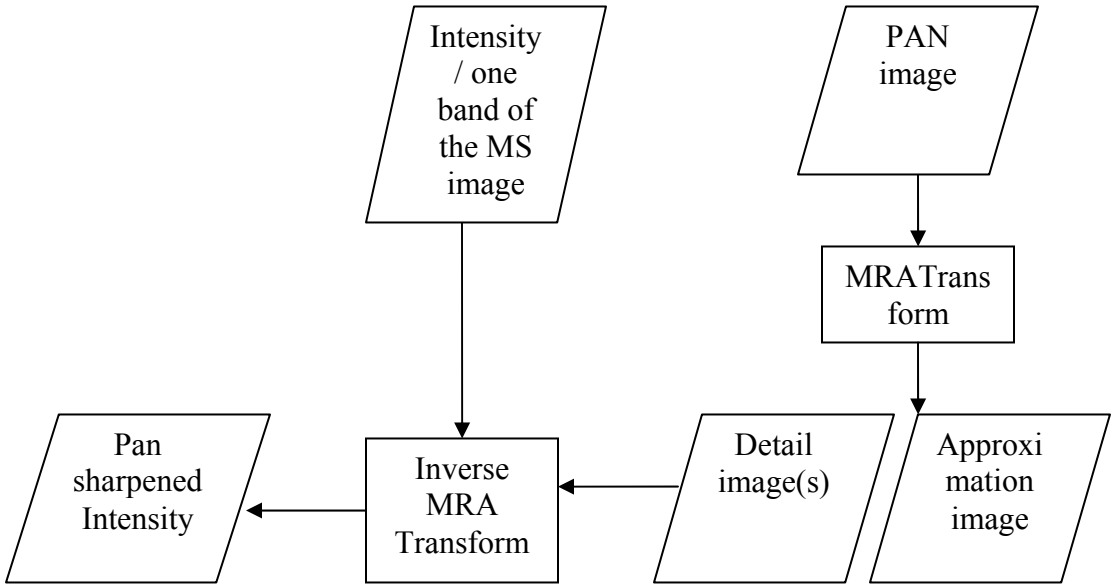


Figure 2-9 Additive MRA based pan sharpening method

www.manaraa.com

A disadvantage of the additive method is the inconvenience caused when the images to be merged do not have a resolution ratio that is a power of two. For example, the resolution ratio between a Landsat 7 MS band at 30 m and SPOT P band at 10 m is 3. A single decomposition of the PAN band would result in an approximation image at 20 m while the MS band has a resolution of 30 m. This would mean inserting an image at resolution 30 m in place of a 20 m image. Even if more than one decomposition is performed the scales of the lowest approximation image of the PAN and the MS image will be different. After two decomposition levels the approximation image of the PAN will have a resolution of 40 m. However, precisely this approach is used by the authors with 2 or 3 decomposition levels. I believe that even though this may have given good results it violates logic as an image with a different resolution has been substituted in place of the approximation image of the PAN.

A workaround this problem is to use an M-scale wavelet transform or a pyramid scheme that decomposes images by a scale other than two, where M is the resolution ratio of the images to be merged. This approach was taken by Shi et al. [37] and Blanc et al. [3], who used a 3-band wavelet transform to merge the TM MS and SPOT P images. Thus the resolution of the approximation image of the PAN after performing one decomposition level will be 30 m and then the 30 m MS image can be substituted in its place. However, very few wavelet software packages are available that provide the general M-band wavelet transforms. This is a big inconvenience for pan sharpening researchers who are mostly experts in geology, hydrology, or some earth science rather than electrical engineers. Moreover, even if a general M-scale transform is written it

www.manaraa.com

would have to be provided for each resolution ratio (e.g. 3, 5, 7, etc.) dataset. Similarly there is also a generalized LP [1] that can decompose images by any scale ratio. Due to time limitations these generalized M-scale transforms could not be implemented. The performance evaluation of these M-scale transforms for pan sharpening is recommended for future work.

Even then the additive merger is applied to images with resolution ratios 2 and 4 in combination with MRA transforms that have M=2 so that the correct scale requirements are met to study its viability and compare it with the substitutive merger.

2.4.5 Substitutive Method

In this technique the MRA is performed on both the PAN and the MS images. Since the aim is to incorporate the details from the PAN image into the MS imagery, the detail coefficients from the MS image and the approximation image of the PAN image are discarded. The approximation image of the MS image is used along with the detail image of the PAN image for synthesizing the pan sharpened image. The method is explained step by step below for sharpening each MS band separately, the same algorithm can be applied to only the Intensity component. Figure 2-10 gives the flowchart of the method for convenience.

1. Both the images are coregistered and the MS imagery is resampled to the same spatial resolution as the PAN imagery.
2. The PAN image is histogram matched to each MS band (as described in the diagram this step is optional). If the histogram matching is performed there are three copies of the PAN image corresponding to each MS image – A_i.

www.manaraa.com

3. The MS and PAN images are decomposed into their approximation and detail images:

$$\{a_{Bi}, d_{Bi}\} = MT(B_i^*), \tag{27}$$

$$\{a_{Ai}, d_{Ai}\} = MT(A_i), \tag{28}$$

4. The detail images of the MS image are replaced by those of the PAN image and the inverse MRA transform is computed to obtain the pan sharpened imagery:

$$a_{Fi} = a_{Fi}, \tag{29}$$

$$d_{Fi} = d_{Ai}, \tag{30}$$

$$F_i = IMT(a_{Fi}, d_{Fi}), \tag{31}$$

www.manaraa.com

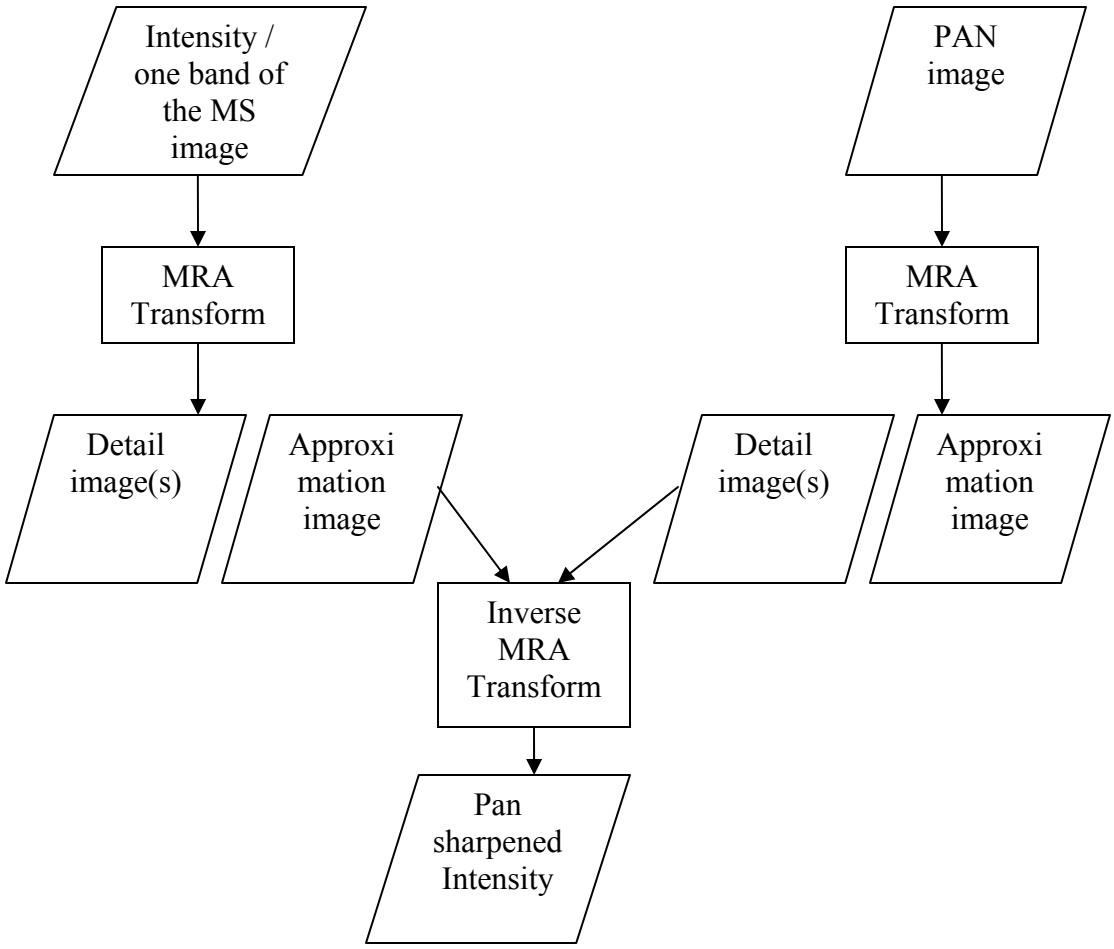


Figure 2-10 Substitutive MRA based pan sharpening method

Many variants of the substitutive method, which better preserve the spectral fidelity of the pan sharpened images with respect to the original MS images, have been suggested over the years. The basic principle behind these techniques is to not discard the detail images of the MS image completely but rather choose the detail coefficients from either the MS or the PAN image at each pixel location in the image based on a selection rule. Some of the popular selection rules found in literature are described in the following subsections. All the selection rules aim to choose the coefficient from either image which

www.manaraa.com

has a higher energy. The detail coefficients with higher magnitude or energy correspond to sharp or drastic changes in the image intensity and thus represent the "salient" features (for e.g. edges) in the image [20]. Since the goal of the pan sharpening is to enhance the edge information in the pan sharpened image, this logic makes sense. The logic behind these rules is that if the MS detail coefficients have the intensity to define the spatial quality they should be retained. The detail coefficients from the PAN image should be injected into the fused image only when required, thus because of this conditional injection of details the change in the radiometry of the MS image is minimized. The various selection rules differ in the way in which "energy" or "salience" is defined.

2.4.5.1 Maximum Amplitude Selection

In the Maximum Amplitude Selection rule (MAS) the coefficients at each pixel given by the following equation [18] –

$$d_F(\vec{p}) = \max\{abs(d_A(\vec{p})), abs(d_B(\vec{p}))\}, \tag{32}$$

2.4.5.2 Window Based Salience

This Window Based Salience rule (WBS) was proposed in [5], Burt et al. reason that since the salient features in the image are generally larger than one pixel, the selection rule must also be applied on more than one coefficient at a time. The salience at a pixel is computed as the variance or energy over a $m \times m$ window centered over it (m is usually 3 or 5) as given below –

www.manaraa.com

$$S_A(i,j,k,l) = \sum_{x=-m}^{+m} \sum_{y=-m}^{+m} d_A(i+x, j+y, k, l)^2 , \quad (33)$$

$$S_B(i,j,k,l) = \sum_{x=-m}^{+m} \sum_{y=-m}^{+m} d_B(i+x, j+y, k, l)^2 , \quad (34)$$

After measuring the salience of each coefficient in the image, the source coefficients are combined using either "selection" or "averaging", based on a match measure between them. The match measure M_{AB} is given by the normalized correlation computed between the two coefficients from A and B around the same neighborhood as used to compute the salience:

$$M_{AB}(i,j,k,l) = \frac{2\sum\sum d_A d_B}{S_A + S_B}, \quad (35)$$

This quantity is a measure of how similar the two images are at a particular pixel location. If the match measure is near +1 it means the images are very similar locally while values near -1 indicate that they are quite different. If the match measure M_{AB} exceeds a threshold α the source coefficients are combined through a weighted average:

$$d_F(\vec{p}) = w_A(\vec{p}) d_A(\vec{p}) + w_B(\vec{p}) d_B(\vec{p}), \quad (36)$$

If M_{AB} is below the threshold, either one of the source coefficients is selected based on the MAS criterion as given by equation $d_F(\vec{p}) = \max\{abs(d_A(\vec{p})), abs(d_B(\vec{p}))\}$, (32). The logic behind selecting the coefficient with higher salience when the match measure is below a certain threshold is to retain a higher contrast in the fused image. Generally, regions with opposite contrast are

www.manaraa.com

associated with a low match measure. For example, if the MS image has a high contrast while the PAN image has a low contrast, and if they are averaged the overall contrast or salience of the fused coefficient is reduced, resulting in blurry looking images.

$$S_A > S_B \Rightarrow d_F(\vec{p}) = d_A(\vec{p}), \tag{37}$$

$$S_B > S_A \Rightarrow d_F(\vec{p}) = d_B(\vec{p}), \tag{38}$$

The weights are computed as follows based on the match measure and the threshold:

$$w_{\min} = \frac{1}{2} - \frac{1}{2}\left(\frac{1 - M_{AB}}{1 - \alpha}\right), \tag{39}$$

$$w_{\max} = 1 - w_{\min}, \tag{40}$$

The higher weight $w_{\max}$ is given to the coefficient with the higher energy and the lower weight $w_{\min}$ is assigned to the other coefficient:

$$S_A > S_B \Rightarrow \{w_A = w_{\max}, w_B = w_{\min}\}, \tag{41}$$

$$S_B > S_A \Rightarrow \{w_B = w_{\max}, w_A = w_{\min}\}, \tag{42}$$

The equations for calculating the weights show that when the match measure exceeds the threshold, $w_{\min}$ is extremely low while $w_{\max}$ is high. As M_{AB} increases towards one the weights become nearly equal. Burt et al. suggested using a quite high threshold ($\alpha = 0.85$); however they also noted that change in the neighborhood size or threshold did not affect the results greatly. In this work a value of 0.75 was chosen for the threshold.

www.manaraa.com

2.4.5.3 Window Based Verification

Li et al. [20] stated that since the window based salience rule can be considered as the cascade of a nonlinear filter with a linear filter, it lacks a clear physical meaning. The variance computation is considered as a nonlinear filter while the wavelet transform is a linear filter. They describe the Window Based Verification rule (WBV), which is a linear filtering operation. In this rule, the maximum amplitude of any pixel in the $m \times m$ window is chosen as the activity or salience of the center pixel under consideration since it indicates the presence of a dominant feature in the neighborhood. Thus, two intermediate images corresponding to the two images to be merged are created whose values are the salience measure of each pixel. These intermediate images are subject to a consistency verification rule which is as follows: if the central pixel comes from image A while the majority of the pixels in the neighborhood are from image B, the central pixel is assigned the value of image B at that location, otherwise the value of image A is retained. The same window size is used when applying the consistency verification rule as that when computing the activity.

2.4.5.4 Addition

Yocky proposed the addition rule (ADD), in which the detail coefficients from both the images are added to generate the detail coefficients of the fused image [47]:

$$d_F(\vec{p}) = d_A(\vec{p}) + d_B(\vec{p}), \tag{43}$$

This way, all the information from the MS image is retained in the fused image. This statement is true in the context of spectral fidelity as it leads to pan sharpened

www.manaraa.com

images that are spectrally closer to the original MS images, since all the MS coefficients are retained. However, in the spatial context this statement is false, as noted by Yocky himself. He suggests this discrepancy may be occurring due to the blockiness of the MS pixels that are preserved in the pan sharpened image. This selection rule can be logically considered identical to the additive method suggested by Nunez et al. [26]. However, experimentally, it is found that numerically there is significant difference in the results obtained from them.

2.4.5.5 Other Possible Rules

Many other rules are possible to combine the coefficients of both the detail images like the ones suggested in [47]. These include –

- Boolean operators (e.g. AND, OR, XOR),
- Maximum value selection rule,
- Conservation of energy, contrast stretching, maximum likelihood ratioing, edge enhancement.

These rules are not studied in this work since the variations due to some of them are minor and in some cases the quality of the fused image can degrade. For example, the maximum value selection rule resulted in poor spatial quality. In reference to the selection rules that were described above, the substitutive method given by equation $d_{Fi} = d_{Ai}$, (30), which discards the detail coefficients of the MS image, will be referred to as the NULL rule since it is equivalent to not applying any selection rule. The NULL and the MAS rules are the most popular in pan sharpening while the WBS and the

www.manaraa.com

WBV rules are more popular in general image fusion applications such as fusion of visible and thermal images, images with different regions out of focus, etc. One of the goals of this research is to compare these rules – MAS, WBS, WBV, and the ADD rules.

2.4.5.6 Computational Complexity of the Selection Rules

The computational complexity of the selection rules described above is also calculated. Compared to the NULL rule the other rules have a significant computational complexity. The simplest substitutive method given by the NULL rule has unit computational cost (the assignment operator), while the MAS rule requires one operation per pixel of the detail images for comparing the magnitudes and thus its computational complexity is N^2. In general, if there are n detail images for the MRA transform used (e.g. n=3 for RWT, n=1 for AWT) and k levels of the transform are computed, the total computational complexity for this rule would be $k \cdot n \cdot N^2$. The ADD rule has the same computational complexity as the MAS rule since it requires one addition operation per pixel. For the window based rules proposed by Burt and Li the number of operations required to compute the salience of each pixel in both the images are at least m^2 and thus the computational complexity increases further – $k \cdot n \cdot m^2 \cdot N^2$. This means the WBV and WBS rules are m^2 times computationally intensive compared to the MAS rule. Since typical values for m are 3 or 5, this rule increases the computation by at least a factor of nine compared to the MAS or ADD rules. Generally, the window based rules are computationally more complex compared to the pixel wise rules.

www.manaraa.com

2.4.6 Model Based Method

2.4.6.1 Introduction

The model based methods can be described as a variant of either the additive or the substitutive fusion methods depending on whether the MRA is performed on both the PAN and the MS images or only the PAN image. There are primarily two models that were found in literature – the AABP model (named after its authors – Aiazzi, Alparone, Baronti, and Pippi) [1] and the RWM model (Ranchin, Wald, and Mangolini) [30]. The basic principle behind these models is that the detail coefficients from the PAN image are inserted into the MS image depending on the local correlation between them and modified prior to insertion to take into account the differences in gray level values of the MS and PAN images.

2.4.6.2 AABP Model

The model of Aiazzi et al. is a conditional fusion scheme where the detail coefficients from the PAN image are inserted into the MS image only when the local correlation between the images exceeds a certain threshold. The logic behind this scheme is that the details from the PAN image at each pixel location in the image must be inserted into the MS image only if the images are similar in nature at that location in the image. This conditional injection of details seems to be proposed to avoid the problem of edges and features getting blurred when the PAN image and the MS image have opposite signatures (i.e., contrast reversal). The authors propose two variants of their scheme one additive (based on the LP) and the other substitutive (based on the RWT). I will refer to

www.manaraa.com

them as the AABP-SUB and the AABP-ADD model respectively. It is not clear why the authors recommended a particular transform for each sub model since it is known that any MRA transform can be used for pan sharpening, (i.e., the AABP-SUB model could be implemented with the LP too). Maybe the authors found through empirical evaluations that these two particular combinations gave better results. This seems to be concurrent with my findings that the additive method based on the LP gives better results than those based on the RWT or the AWT. The additive scheme is described in detail below since the authors mention they had slightly better results with it:

1. The PAN image is decomposed using the LP,
2. The MS image is resampled to the same resolution as the PAN image using a 23-tap polynomial filter. It must be noted that here the choice of filter to upsample the MS images should not be critical, a bilinear or bicubic interpolation should be adequate,
3. A Local Correlation Coefficient (LCC) map is calculated between the resampled MS image and the approximation image of the PAN. The LCC is just the correlation coefficient but computed over a *mxm* sliding window for each pixel in the entire image,
4. At each pixel location the LCC is matched against a threshold θ, if it exceeds the threshold the detail coefficient at that position is inserted into the MS image. Otherwise no injection of detail coefficients takes place:

www.manaraa.com

$$LCC(\vec{p}) > \theta \Rightarrow d_F(\vec{p}) = d_A(\vec{p}), \tag{44}$$

$$LCC(\vec{p}) \leq \theta \Rightarrow d_F(\vec{p}) = 0, \tag{45}$$

5. Before inserting the detail coefficients, they are multiplied by a Local Gain (LG) factor, which is the ratio of the standard deviation of the MS image to the standard deviation of the approximation image (corresponding to PAN):

$$LG(i, j) = \frac{\sigma_B(i, j)}{\sigma_A(i, j)}, \tag{46}$$

   This is done in order to scale the values of the detail coefficients of the PAN image to that of the MS image,

6. The pan sharpened image is synthesized by adding the resampled MS image with the scaled detail image of the PAN. This step is repeated for each MS band.

In the substitutive scheme the MRA is performed on both the PAN and the resampled MS image using the RWT. The LCC is computed between the approximation images of the PAN and MS images. The thresholding and the scaling steps are identical to that of the additive scheme. In case the LCC is below the threshold, the MS detail coefficients are retained. In the AABP model, two parameters can be varied – the threshold θ and the size of the window m. The authors studied the performance of the model with respect to these parameters and found that they have little effect on the quality of the pan sharpened images. The quality slightly improved with a larger window, and the threshold range of (-1, 0.8) gave almost similar results. A threshold higher than 0.8 resulted in poor quality pan sharpened images. This effect is understandable as there

www.manaraa.com

will be fewer and fewer coefficients that will have such a high correlation (otherwise there would be no need for pan sharpening), and consequently details will not be injected into the MS image.

2.4.6.3 RWM Model

The RWM model is also a conditional fusion scheme like the AABP model. However, unlike the AABP model in which the LCC is computed between the approximation images of the PAN and MS in case of the AABP-SUB method (or the resampled MS and approximation of the PAN in case of the AABP-ADD method), in the RWM model it is computed between the detail images itself. Since the authors of [30] found that both models gave almost identical results, the RWM model was not implemented in this work.

2.4.7 Summary

This chapter gave an overview of many pan sharpening techniques. I have shown that pan sharpening techniques can be divided into two categories – component substitution (COS) based and arithmetic combination based methods (AMC). The advantages and disadvantages of each method are also given. The HPF technique laid the foundation for the MRA based methods. Since the focus of this research is the MRA techniques a thorough review of the MRA techniques is also given. MRA based methods preserve the spectral fidelity better than other methods because they extract only the high frequency information from the PAN image to inject into the MS image.

www.manaraa.com

CHAPTER III

RESEARCH OBJECTIVES AND METHODOLOGIES

3.1 Introduction

This chapter first outlines the research objectives and then explains the methods and metrics used for quantitatively assessing the quality of the pan sharpened images. This chapter describes the details of the study – how the experiments were conducted and what the criteria and metrics used to decide which algorithm is best were.

3.2 Research Objectives

Chapter 2 gave a comprehensive overview of the various MRA based pan sharpening techniques. As stated previously, these variations arise out of the choice of preprocessing options, the MRA transform used, and the methods used to synthesize the coefficients of the pan sharpened image. This research studies these variations, the effect of the preprocessing options, the MRA transform, etc. in order to make recommendations regarding which variation gives the highest spectral fidelity. Only options or parameters that seem to be appropriate to this study are addressed. In this case the one deemed to be the most appropriate is chosen based on the literature review. The different parameters and preprocessing options in the MRA based techniques have a small or significant effect

www.manaraa.com

on the spectral and spatial quality of the pan sharpened image. For example, the mother wavelet used in the RWT based pan sharpening can have a subtle or significant effect on the quality of the pan sharpened image. A major contribution of this work is that I have used multiple datasets from different sensors before arriving at my conclusions. Sometimes researchers test their algorithms on a single dataset of small size (512x512 pixels) to arrive at their conclusions.

The following subsections describe the parameters that were investigated and the reasoning for the approach adopted.

3.2.1 Effect of Radiometric Normalization

The importance of the radiometric normalization step was given in section 2.4.3.1, and most authors prefer to implement it. The effect of this preprocessing step on the quality of the pan sharpened images will be analyzed by simply turning this option on and off. Since in this study each MS band is merged individually the PAN band is histogram matched to each MS band and thus there are N modified PAN bands for the N multispectral bands to be merged.

3.2.2 MS Transformation

Section 2.4.3.2 mentions that there can be loss of detail information and artifacts in the pan sharpened image if there are contrast reversals in the scene and if special adaptive processing is not applied in these regions for each band. Otherwise, if contrast reversal is not a concern (for e.g. if such regions are scarce) the MS bands can be transformed into IHS or PC before applying the MRA based pan sharpening.

www.manaraa.com

In this study, due to time limitations, adaptive processing for contrast reversal was not implemented; moreover most of the datasets do not exhibit contrast reversal. Only one dataset (tm-vorarlberg) exhibited contrast reversal, and it was found that in this case avoiding the radiometric normalization step gave fairly good results. Thus I could have applied a IHS or PC transformation before performing the pan sharpening. However, it is chosen not to do so because the idea is that in future the results from this work will be used and adaptive processing for contrast reversal must be applied.

3.2.3 Choice of the Wavelet Basis

Some researchers have concluded that the choice of the mother wavelet does not matter [32]. Some claim that the higher order, smoother wavelets gave better results [18] while others claim that the shorter wavelets gave better results [44]. However, there are not many details in the literature that explain why a certain wavelet basis was better than the others or the results were based on different metrics. This work will investigate the effect of the wavelet basis on the pan sharpened image in more detail using quantitative metrics.

Two popular wavelet families will be investigated – the Daubechies wavelets and the biorthogonal wavelets. The Daubechies wavelets are nonlinear phase while the biorthogonal wavelets are linear phase [4]. Daubechies wavelets with different number of vanishing moments were used to observe the effect of the number of vanishing moments on the results. As the number of vanishing moments increase, the wavelet coefficients give a better approximation of the signal [4]. In this dissertation I use the notation ‘dbN’ to denote a Daubechies wavelet having N vanishing moments (e..g ‘db2’ denotes a

www.manaraa.com

Daubechies wavelet with two vanishing moments). The experiments were limited to the Daubechies wavelets – 'db1', 'db2', 'db5', 'db10', and 'db20'. These wavelets were studied to see the effect of very short, medium and long wavelets on quality of the pan sharpened images. In order to observe only the effect of the wavelet, no selection rules on the high frequency coefficients are applied.

The Daubechies wavelets belong to a class of filters called the nonlinear phase filters. The choice of these filters in image compression leads to higher artifacts and less coding gain compared to another class of filters called the 'linear phase filters'. There were two purposes for performing these set of experiments – one is to find the mother wavelet that results in the highest spectral fidelity and the second was to observe if the linear phase filters had any advantage over the nonlinear filters. All the biorthogonal filters that were available in the MATLAB toolbox were investigated: bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8.

Similarly, the filters used for the LP based pan sharpening will be studied in order to determine the one which gave the highest spectral fidelity. There are many filters that can be applied to the LP (e.g. the binomial filters, Quadrature Mirror Filters (QMF), and the Gaussian filters). Literature shows that for LP or AWT based image fusion the Gaussian filters are the most popular. These filters will be used to find an empirically best filter for LP based pan sharpening. Since the AWT is just an oversampled version of the LP the same filters will be used for the AWT too to find the one which gives the highest spectral fidelity.

www.manaraa.com

3.2.4 Comparison of Different Coefficient Synthesis Methods

Chapter 2 explains that there are two main methods in the MRA based technique to merge the MS and PAN images – the additive and the substitutive method. The additive method based on the RWT, AWT, and the LP will be compared to the substitutive method based on the RWT. The comparison is restricted to only the RWT based substitutive method since a detailed investigation of the substitutive methods showed that this approach best preserves the spectral fidelity compared to the AWT or the LP.

3.2.5 Choice of the Selection Rule

Five main selection rules for the substitutive method were described in section 2.4.5 – NULL, MAS, WBS, WBV, and ADD. From the logical explanation of these rules it is clear that the advanced rules will increase the spectral fidelity compared to the NULL rule. This work will compare these rules and ascertain which of them gives the highest spectral fidelity. The comparison was restricted to the rules – NULL, MAS, WBS, and WBV as they are the most popular in image fusion. The ADD rule resulted in poor spatial quality and thus, the results for this rule are not compared to the NULL rule.

In pan sharpening most of the researchers employ either the NULL [21], [52] or the MAS rule [19], [46]. Very few articles were found that used the window based rules [20] for pan sharpening, which seem to be highly favored in image fusion research. To the best of my knowledge a comparison of the different selection rules has not been undertaken in the context of pan sharpening. These rules have been compared in the context of image fusion by few authors [20], [51]. The finding of [20] was that the WBV

www.manaraa.com

rule outperformed MAS and WBS; while [51] provides a broader conclusion that the window based methods perform better than pixel based methods for choosing the coefficient from either image. It cannot be assumed that the same results apply to pan sharpening images because not only are the input image characteristics different, but the aim of the fusion and metrics used to assess the quality of the fused images are also different. In general image fusion (which was done by the authors of [20], [51]) the images to be fused were out of focus images. Such images are simulated by taking an image and blurring it using some type of lowpass filter or introducing some kind of distortion (noise) in different regions. For example an image C is blurred in the left half and right half respectively creating two images A and B. These images A and B are then fused to create an image F which must be as close to C as possible. The primary metric used by many researchers, including the above authors, is the Root Mean Square Error (RMSE) between the original image C and the fused image F. Such a reference image is not available in pan sharpening.

The goal of pan sharpening is to create a high resolution MS image which best preserves the spectral similarity with respect to the original MS image while incorporating the spatial resolution of the PAN image. Thus, the selection rules will be judged by two types of metrics – spectral quality metrics and spatial quality metrics, (i.e., it should increase the spectral fidelity while retaining the spatial resolution of the PAN image). These metrics are described later in this chapter.

www.manaraa.com

3.2.6 Choice of the MRA Transform

In the additive or substitutive method any MRA transform can be used, it was stated in section 2.4.3.3 states that the most popular ones are the DWT, RWT, AWT and the LP. The number of articles found during the literature survey that used the RWT are far more than those that used the AWT, LP or the DWT.

An advantage of the transforms like RWT, AWT, and the LP is that they are shift invariant (the LP is approximately shift invariant) while the DWT is shift variant. Many authors have noted that the shift variance of the DWT causes artifacts in the fused images [14], [15], [18], and [51]. Thus the DWT images are not analyzed further. These artifacts can be minimized by using linear phase or biorthogonal filters [15], [51] but they cannot be entirely removed. The DWT pan sharpened images produced in this work had artifacts, surprisingly for the shorter biorthogonal filters the artifacts worsened. In general, the longer Daubechies and biorthogonal wavelets produced fewer artifacts.

Comparison of the different MRA transforms has been made by various authors but very few authors have compared them in the context of pan sharpening. A comparison between the LP and the RWT in the context of pan sharpening was done by Aiazzi et al. [2] and Ranchin et al. [30]. Aiazzi et al. [2] proposed the AABP model which was explained in section 2.4.6.2, this model had two variations. The additive model was based on the LP while the substitutive model was based on the RWT. The authors applied both the variations to pan sharpen SPOT images and found that they gave almost similar results. Interestingly in another joint study [30], they found that the LP based AABP model outperforms the RWT based model proposed by Ranchin et al.

www.manaraa.com

In both the studies, the authors assess the spectral quality of the pan sharpened images by comparing these images to the ideal high resolution MS image that would have been acquired by a MS sensor having the same resolution as the PAN image. This comparison requires the presence of a high resolution MS image which cannot be acquired in practice, thus the original MS-PAN images are degraded so that the MS image has a lower resolution and the PAN image has the same resolution as the original MS image. After pan sharpening the degraded MS image it has the same resolution as the original MS image and thus it can be compared with it [44]. Wald et al. [44] believe the results can be extrapolated to higher resolution, i.e. the quality of the pan sharpened image produced by merging the original MS-PAN will be of the same quality. In this work I prefer to use a different approach – measure the spectral distortion in the pan sharpened image by comparing it with the original MS image. Our studies show that the latter approach should be preferred. These issues will be explained in detail in sections 3.3.3.4, 3.3.3.5 and 3.3.4.

Thus in this work I propose to compare the RWT, AWT, and the LP pan sharpened images using the substitutive method in order to see which one gives the highest spectral fidelity. Since the results of each transform are also affected by the wavelet basis used (RWT) or the filter used (AWT, LP) first the wavelet (or filter) which gives the best results for each transform will be determined and used for comparison with the other transforms. The remaining parameters of the fusion process should be kept identical – the radiometric normalization (applied for each of them or none of them) and the number of decomposition levels of the transform. The quality of the pan sharpened

www.manaraa.com

images will be measured by directly measuring the spectral distortion between the MS and the pan sharpened images.

3.2.7 Effect of Directional Selectivity of the Transform

The directional selectivity of a transform can be defined as its ability to capture the details of an image in different spatial directions. The details generated by the LP or the AWT do not have a specific spatial direction, thus it is not very effective for pattern recognition of textures [23]. On the other hand the detail images generated by the wavelet transforms – DWT and the RWT have a spatial direction. These transforms produce three detail images at each scale which capture the image details in three directions – horizontal, vertical and diagonal. This improved directional selectivity is helpful in efficient texture discrimination.

It is generally believed that transforms with a higher directional selectivity give better results for image fusion. Li et al. [20] found that the DWT performed better than the LP for fusion of out of focus images and reasoned that this was because of the higher directional selectivity of the DWT.

The goal of this research is to determine if higher directional selectivity improves the quality of the pan sharpened images. This comparison can easily be done by comparing the results of the LP or AWT and comparing them with the results of the RWT. However, the differences between the RWT and the AWT are not only different directional selectivity but also a different multiresolution operator (wavelet or filter), and a different subsampling rate in the case of the LP. Since these two factors also have an

www.manaraa.com

effect on the quality of the pan sharpened images the effect of the directional selectivity cannot be studied in isolation.

Thus, I plan to use the Steerable Pyramid Transform (SPT) [38] to pan sharpen the images. Since the SPT can have any desired directional selectivity the same MS-PAN image pair will be pan sharpened using the SPT but with a different number of detail images each time. The spectral quality of the pan sharpened images created using a different number of orientation bands will be analyzed.

3.2.8 Estimation of the Minimum Number of Decomposition Levels

In the substitutive method described in section 2.4.5, both the MS and the PAN images are decomposed 'L' times (where generally L is chosen to be between one and four). The spectral and spatial qualities of the pan sharpened images are affected by the number of decomposition levels. Rockinger [32] concluded that the presence of larger objects required more levels of decomposition for a good fusion, while Nunez et al. [26] report that 2 to 3 levels of decomposition are sufficient. Preliminary studies have shown that the quality of the pan sharpened images produced by the MRA based technique is a function of the number of decomposition levels and the resolution ratio of the PAN-MS pair. If fewer levels of decomposition are applied, the spatial quality of the pan sharpened images is less satisfactory, while the spectral similarity between the original MS and pan sharpened images decreases if excessive levels are applied. Thus, this work attempts to determine the best or the minimum possible number of decomposition levels for the wavelet based fusion, i.e., which yields sufficient spatial quality and the highest spectral

www.manaraa.com

fidelity. The sufficient spatial quality implies that the pan sharpened image has the spatial resolution of the PAN image.

In practice, remote sensing images are very large. They may have tens of thousands of rows and columns in each band. Thus, each extra decomposition level results in a longer wait time for the result. This work investigates the quality of the pan-sharpened images produced by using various levels of decomposition and by merging PAN-MS image pairs with different resolution ratios.

The aim of this study is to make recommendations on the best possible number of decomposition levels required for merging images with a particular resolution ratio. For this purpose MS-PAN images with four different resolution ratios are used – 2, 3, 4, and 5, since pan sharpening images with ratios 2, 3, and 4 is common. The resolution ratio 5 is chosen to add more breadth to the study. For these experiments it was decided to use Wald's property 1 and 2 both.

The results are expected to act as a guideline for practitioners who desire to implement the MRA based fusion scheme. These experiments were only carried out for the RWT scheme, but the results should also apply to the AWT and LP methods. To the best of my knowledge, a detailed investigation of the effect of the number of decomposition levels on the fusion results of PAN and MS imagery has not been conducted.

It was also found that Zhou et al. [52] made somewhat similar observations. They merged a SPOT PAN with LANDSAT TM images using 2 and 3 decomposition levels and found that the pan sharpened image created with 2 decomposition levels had higher

www.manaraa.com

spectral quality while the pan sharpened image created with 3 decomposition levels had higher spatial quality as measured by their spatial quality metric – the High Pass Correlation Coefficient (HPCC), which is described below. However, they just worked on one resolution ratio (3) and my conclusions on the number of decomposition levels also differ from theirs.

3.3 Image Quality Metrics

The emphasis of this work is to produce pan sharpened images that retain the spectral fidelity of the original MS images while achieving a high spatial quality. A quantitative assessment of the spectral, as well as spatial quality of the pan sharpened image, must be done. This section describes the metrics used to assess the pan sharpened images. The findings and conclusions of this research will be based on these metrics. There are mainly two ways to evaluate the quality of the fused images or any image in general – subjective and objective metrics [16].

3.3.1 Subjective Metrics

Subjective metrics are based on the opinion of human observers. As an example, consider a researcher who must decide if a new compression algorithm is better than an older one. The compressed images produced by both techniques and the original image are presented to a group of observers. Each observer ranks the image on a given scale and the results are averaged. However, this technique is questionable since individual opinions vary considerably. Moreover, it may be a costly metric since it is difficult to

www.manaraa.com

find volunteers, in which case the observers must be paid. Thus, these metrics should be kept to a minimum.

3.3.2 Objective Metrics

Objective metrics on the other hand are quantifiable. They exploit the pixel level differences in the images, correlation between images and the gray level distributions (histogram) [43] to assess the image quality. The pan sharpened image must not only have all the spatial details of the PAN image, but also be spectrally similar to the original MS image as possible. The following sections describe the objective spectral and spatial quality metrics.

3.3.3 Spectral Quality Metrics

This section describes the various spectral quality metrics used to evaluate the spectral fidelity of the pan sharpened images with respect to the original MS images are described. Since the goal is to preserve the radiometry of the original MS images, any metric used must measure the amount of change in DN values in the pan sharpened image compared to the original image. The following subsections define a few spectral quality metrics.

3.3.3.1 Correlation Coefficient

The correlation coefficient measures the closeness or similarity between two images [43]. It can vary between –1 to +1. A value close to +1 indicates that the two images are very similar, while a value close to –1 indicates that they are highly dissimilar. The formula to compute the correlation between two images A and B, both of

www.manaraa.com

size NXN pixels is given by:

$$Corr(A \mid B) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}(A_{i,j} - \bar{A})(B_{i,j} - \bar{B})}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}(A_{i,j} - \bar{A})^2 \sum_{i=1}^{N}\sum_{j=1}^{N}(B_{i,j} - \bar{B})^2}}, \tag{47}$$

Various correlation coefficients are computed to evaluate the spectral quality of the pan sharpened images [43]. The inter-correlation between each pair of the unsharpened bands and the sharpened bands was computed and compared. For example, $Corr(B_1, B_2)$ is the interband correlation between bands 1 and 2 before fusion, and $Corr(F_1, F_2)$ is the interband-correlation after fusion. Ideally a zero change in the correlation values would be desirable, i.e. if $Corr(B_1, B_2)$ was 0.94, the ideal value for $Corr(F_1, F_2)$ would be 0.94. Thus, if there are three bands being fused the inter-correlation between each pair of the three unsharpened bands and the three sharpened bands was computed and compared:

$$Corr(B_1 \mid B_2), Corr(F_1 \mid F_2), \tag{48}$$

$$Corr(B_2 \mid B_3), Corr(F_2 \mid F_3), \tag{49}$$

$$Corr(B_1 \mid B_3), Corr(F_1 \mid F_3), \tag{50}$$

Then, the correlation between each sharpened and unsharpened band $Corr(B_i^*, F_i)$ (for i = 1, 2 … N if there are N bands) is computed. The ideal value for this is 1. Since the pan sharpened images are larger (more pixels) than the original MS image 'B' it is not possible to compute the correlation or apply any other mathematical operation between them. Thus, the resampled MS image B^* is used for this comparison.

www.manaraa.com

This is acceptable since the resampling produces little change in the radiometry of the original images. Similarly, the correlation between the resampled MS image and the PAN image $Corr(B_i^* \mid A)$ is computed and it is compared with the correlation between the fused and PAN images $Corr(F_i \mid A)$. Ideally, the correlation between the fused and PAN image should be the same as that between the original MS and PAN image.

3.3.3.2 Root Mean Square Error

The Root Mean Square Error (RMSE) between each unsharpened MS band and the corresponding sharpened band is also computed as a measure of spectral fidelity [21]. It measures the amount of change per pixel due to the processing (e.g. pan sharpening) and is described by:

$$RMSE_k = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\sqrt{(B_k^*(i,j) - F_k(i,j))^2}}{N^2}, \tag{51}$$

During this work, it was found that the RMSE has a higher resolution compared to the correlation coefficient. This statement means that if the performance of the two algorithms is almost identical to each other, then the RMSE can better distinguish which one is better. For example, if the pan sharpened images produced by algorithms 1 and 2 have a correlation coefficient of 0.99 with respect to the MS image, it means the spectral quality of both algorithms is identical. On the other hand, if the RMSE values for the two corresponding images are 2.34 and 2.12 respectively, clearly algorithm 2 results in a higher spectral quality compared to algorithm 1, and only the RMSE can clarify this distinction.

www.manaraa.com

3.3.3.3 Histogram

The histograms of the original MS and the pan sharpened bands must be evaluated [43]. If the spectral information has been preserved in the pan sharpened image, its histogram will closely resemble the histogram of the original image. Though the histogram is a quantitative measurement, the comparison of two histograms can be carried out either subjectively (visually) [43], [44] or objectively by computing the shift in the mean of the original MS image [27]. The shift in the mean value of a band i is computed as:

$$SM_i = \frac{\mu(F_i) - \mu(B_i)}{\mu(B_i)} \times 100, \tag{52}$$

The symbol $\mu(X)$ denotes the mean of an image X. A fusion technique that preserves the spectral fidelity of the images should result in low values for SM (i.e., ideally the mean of the image should not change).

3.3.3.4 Wald's Property 1

Wald et al. [44] arrived at three properties that a sharpened image must have in order to assess its spectral quality. They state the first property as follows "Any synthetic image F(h) once degraded to its original resolution *l*, should be as identical as possible to the original image B(l)". The synthetic image refers to the pan sharpened image in this context, while *h* and *l* are the resolution of the PAN and the MS images respectively.

This property is tested by using a low pass filter to spatially degrade the sharpened image to its original low resolution. Wald et al. noted that the choice of the filter is not very significant . In this study, a bicubic resampling was applied to obtain the

www.manaraa.com

degraded image F'(l) from the sharpened images. One could also apply a wavelet transform and compare the approximation image obtained from the pan sharpened image and compare it with the original MS image. Nevertheless, the main point is that the same degradation method must be used to compare pan sharpened images produced by different algorithms. Statistical comparisons are then done between the two images by using various metrics like SM, correlation between each band i of F'(l) and B(l) ($Corr(B(l) \mid F'(l)_i$), the difference in variances of the two images relative to the variance of the original MS image, and a few other quantities. No other literature articles were found that calculated property 1 as a measure of the spectral fidelity of the images.

3.3.3.5 Wald's Property 2

The second property states that "Any synthetic image F(h) should be as identical as possible to the image B(h) that the corresponding sensor would observe with the highest resolution *h*". However, testing the second property requires the existence of a higher resolution image captured by the MS sensor that has the same resolution as the PAN image. Such an image does not exist (otherwise there would be no need for pan sharpening). Wald et al. propose the following approach to circumvent this difficulty:

1. The original PAN A(h) and MS images B(l) are degraded to a lower resolution. The PAN image is degraded to have the MS image's resolution *l* and the MS image degraded to have a resolution *s*. The resolution ratio *s/l* must be maintained the same as *l/h* in order to make correct comparisons. The degradation is done by

www.manaraa.com

lowpass filtering them over a nxn neighborhood and downsampling the filtered images by n along both rows and columns (n = *l/h* the desired resolution ratio)

2. A(l) and B(s) are fused to obtain the pan sharpened image F(l) at the same resolution as the original MS image – B(l).
3. The quality of the pan sharpened image F(l) can be compared to B(l) by using various statistical methods like the correlation coefficient, shift in mean, etc. as suggested for property 1.

In this work, in order to measure property 2, I take a slightly different approach to create A(l) and B(s). In Yocky [46], A(l) is created by taking the average or a weighted average of the MS bands. The method to obtain the degraded MS image is the same as Wald's. I believe that in Wald's method if the original PAN and MS images are misregistered this misregistration is retained in the degraded PAN and MS images and thus the results will be affected. In Yocky's method the degraded MS image will always be coregistered to the PAN image. This is a good technique for someone wanting to research pan sharpening algorithms, but without the skills to coregister images or to a MS-PAN image pair of the same ground location. Moreover, the spectral differences in the PAN and MS images are minimal.

3.3.3.6 Wald's Property 3

The third property is just an extension of the second property. It has been selectively implemented in this work for conciseness. It is stated as follows, "The multispectral set of synthetic images F(h) should be as identical as possible to the multispectral set of images B(h) that the corresponding sensor would observe with the

www.manaraa.com

highest resolution *h*". A comparison of the statements of properties 2 and 3 shows that they are almost identical except for the fact that in the third property the whole set of images is compared rather than a band to band comparison.

The calculation of the third property involves subjective evaluation (i.e., visual analysis). The subjective assessment suggests that the color composites of images F(h) and B(h) must agree visually. The MS and the pan sharpened images are simultaneously displayed on the screen and the subject attempts to observe the differences in contrast, color, etc. between the two images. However, the ability of humans to distinguish minute differences between the colors of the two images is limited as Wald et al. themselves note. Thus, this method would be acceptable if a general idea is required about whether a pan sharpened image has the same colors as the MS image. However, when a decision has to be made to choose from among many pan sharpened images that have minute differences (like those produced by small variants of the MRA technique) the visual analysis clearly has shortcomings. As a result property 3 was not measured in this work.

3.3.4 Conflict of Wald's Property 1 with Property 2

In the following discussion the spectral fidelity metrics that calculate the change in the radiometry of the pan sharpened image with respect to the resampled MS image (e.g. the correlation coefficient or RMSE given in sections 3.3.3.1 and 3.3.3.2 above) will be referred to as property 1 type metrics. Wald's property 1 measures how similar the pan sharpened image is to the original low resolution MS image, while properties 2 and 3 measure how similar the pan sharpened image is to the ideal high resolution MS image. The literature survey showed that researchers are divided into two groups – those that

www.manaraa.com

measure the performance of their algorithms based on property 1 or its variants and those that measure the performance based on property 2 (and 3). For example, Aiazzi et al. [1], Lemeshewsky [18], Wald et al. [44], Schowengerdt [34] prefer the use of property 2 while Li et al. [21], Shi et al. [37] and Yocky [4746] prefer property 1 type metrics.

During this research it was found that if an algorithm produces pan sharpened images that satisfy properties 2 and 3 well, then it will not satisfy property 1 well. Though property 3 is not measured in this research, it is believed that it will follow the pattern of property 2 (i.e. if a pan sharpening technique improves property 2 it will also improve property 3). The conflict of property 1 and 2 is demonstrated by pan sharpening a MS image degraded to a lower resolution and creating a PAN image by averaging the three MS bands (as suggested by Yocky [46]) and comparing it with the original MS image. The geovantage_ratio4 dataset was pan sharpened using five different decomposition levels – 1 to 5. As the decomposition levels are increased, the spatial resolution of the pan sharpened images improves. Table 3-1 shows the property 1 and 2 metrics for the different pan sharpened images, the three values in each cell are the values for MS bands 1, 2, and 3. As the decomposition levels are increased from 1 to 5 there is a steady improvement in property 2 metrics, which means the pan sharpened image is becoming a good approximation of the high resolution MS image. However, the property 1 metrics deteriorate, which means the pan sharpened image is losing its original spectral characteristics. Among all the pan sharpened images, the one produced from five decomposition levels best satisfies property 2, however the property 1 metrics are the lowest for this image.

www.manaraa.com

Table 3-1 Conflict of Wald's property 1 and 2

	Property1		Property 2	
Num. decomposition levels (L)	Correlation	RMSE	Correlation	RMSE
1	0.99 0.98 0.99	8.74 8.16 8.9	0.94 0.93 0.94	18.0 17.0 19.0
2	0.98 0.98 0.98	10.0 9.06 9.97	0.96 0.96 0.96	14.0 14.0 16.0
3	0.97 0.97 0.97	12.2 10.4 11.5	0.98 0.98 0.97	10.0 9.9 12.0
4	0.96 0.96 0.96	14.4 12.1 13.4	0.99 0.99 0.98	8.1 8.0 10.0
5	0.95 0.95 0.96	15.54 13.2 14.4	0.99 0.99 0.98	8.0 7.7 9.7

This experiment demonstrates that maximizing both property 1 and 2 at the same time are conflicting goals. The main reason why MRA techniques are pursued for pan sharpening was to reduce the spectral distortion in the pan sharpened images. The goal of pan sharpening is two-fold – achieving the spatial quality of the PAN image while preserving the spectral quality of the MS images. It was found through experiments that maximizing the property 2 metrics is not really required in order to get good spatial quality. In the above experiment it was found that according to property 2 metrics L=5 is the optimal decomposition level, the spatial quality of the image pan sharpened with L=3 looks sufficient. It is seen that the property 1 metrics of L=3 image are better than those for L=5.

One of the goals of this research was to produce pan sharpened images that would

www.manaraa.com

be used for supervised classification where the spectral signatures used are derived from the MS images and used to classify the pan sharpened image. Any change in the radiometry of the imagery increases the chances of misclassification. In this case, property 1 and its variants are a better measure of the spectral distortion of the pan sharpened images. Moreover, property 1 and its type of metrics are useful in a real world scenario where the ideal high resolution MS image will not be available. Thus in this work the spectral quality is evaluated based on property 1 and its type metrics will be used.

3.3.5 Spatial Quality Metrics

The evaluation of the spatial quality of the pan sharpened images is equally important since the goal is to retain the high spatial resolution of the PAN image. A survey of the pan sharpening literature revealed there were very few papers that evaluated the spatial quality of the pan sharpened imagery. Consequently, there are very few spatial quality metrics found in the literature. It was found that these metrics had some drawbacks and thus, the spatial quality computed by them was in error. This is explained in section 3.3.6.

3.3.5.1 Mean Gradient

Image gradients have been used as a measure of image sharpness [33]. The gradient at any pixel is the derivative of the DN values of neighboring pixels. Generally sharper images have higher gradient values. Thus, any image fusion method should result in increased gradient values because this process makes the images sharper compared to

www.manaraa.com

the low resolution image. The mean gradient defines the contrast between the details variation of pattern on the image and the clarity of the image [37]. The mean gradient $\bar{G}$ of an image X is given by:

$$\bar{G} = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sqrt{\frac{\Delta I_x^2 + \Delta I_y^2}{2}}, \tag{53}$$

$$\Delta I_x = X(i+1, j) - X(i, j), \tag{54}$$

$$\Delta I_y = X(i, j+1) - X(i, j), \tag{55}$$

where ΔI_x and ΔI_y are the horizontal and vertical gradients per pixel. Ryan et al. also used the mean gradient to characterize the image quality for Ikonos images [33]. However, the gradient calculation implemented by them was slightly different, as they used a Roberts operator that measures the gradient in the 45 degree diagonal directions, i.e., ΔI_x and ΔI_y are given by:

$$\Delta I_x = X(i, j) - X(i+1, j+1), \tag{56}$$

$$\Delta I_y = X(i+1, j) - X(i, j+1), \tag{57}$$

3.3.5.2 High Pass Correlation Coefficient

This approach was first proposed by Zhou et al. [52] to measure the amount of edge information from the PAN image is transferred into the fused images. The high spatial resolution information missing in the MS image is present in the high frequencies of the PAN image. The pan sharpening process injects the higher frequencies from the PAN image into the MS image. Thus, they propose that the correlation coefficient

www.manaraa.com

between the highpass filtered PAN and the pan sharpened images would indicate how much spatial information from the PAN image has been incorporated into the MS image. A higher correlation between the two highpass filtered images implies that the spatial information has been retained faithfully. This correlation coefficient is called the Highpass Correlation Coefficient (HPCC). The authors made use of a Laplacian operator as the highpass filter, whose coefficients are given by:

$$HP = \begin{bmatrix} -1 & -1 & -1 \\ -1 & +8 & -1 \\ -1 & -1 & -1 \end{bmatrix}, \tag{58}$$

Though no reason is stated for the preference of this particular filter it is believed that any highpass filter (e.g. Sobel, Prewitt [16]) can be used for this purpose. Let Y(A) be the high pass filtered PAN image and $Y(F_i)$ be the high pass filtered pan sharpened images. Then the HPCC is given by:

$$HPCC_i = Corr(Y(A), Y(F_i)), \tag{59}$$

3.3.6 Limitations of the Spatial Quality Metrics

The goal of the MRA based pan sharpening is to produce pan sharpened images with the highest spectral fidelity while the spatial quality is sufficient such that all the structures observed in the PAN image can be observed easily in the pan sharpened image. Though it is not investigated whether a small increment in the spectral fidelity will result in improved classification, generally the algorithm is tweaked to give as high a spectral fidelity as possible assuming that even the slightest improvement will be beneficial. The same cannot be said about the spatial quality. It is sufficient that the user should be able

www.manaraa.com

to observe all the details of the PAN image in the sharpened image. Of course the images with higher spatial quality look crisp and sharp, but it has been found that such an increase in spatial quality is at the expense of reduced spectral fidelity.

Thus, the task is to define the term "sufficient spatial quality". This is a very difficult task and cannot be measured by any metric (i.e. how can it be said if a HPCC value of 0.94 is sufficient or not?). Perhaps a HPCC of 0.95 is sufficient, but a value of 0.94 is not sufficient, i.e. the sharpened image do not contain all the details observed in the PAN image. Moreover, this value could be different for different images. Whether the pan sharpened image has all the details in the PAN image can only be determined through visual analysis. Thus, visual analysis played an important role in determining the spatial quality of the images. Perhaps this is the reason for so few studies on the spatial quality of the pan sharpened images.

During this study, the above mentioned metrics were measured and researched. A few anomalies were found with all of them. For one of the datasets, the HPCC values for the three bands sharpened with the Brovey method were lower than those of images sharpened from the wavelet based method using two decomposition levels, implying the Brovey method sharpened images contain less high frequency information than those obtained from the wavelet based method. Visual analysis, however, clearly contradicts this result. Moreover, contradictions were also noticed for the images sharpened by the wavelet-based method. For the first dataset, images sharpened using five and nine levels of decomposition had lower HPCC values compared to the images sharpened using two, three and four levels. Visual analysis showed that the image sharpened using two levels

www.manaraa.com

was not sufficiently merged and should have lower HPCC values compared to those for five and nine levels. It is suspected that HPCC is affected not only by the correlation between the edges, but also by the pixel values of the filtered images. Similar anomalies were noted for the mean gradient metric. When the MS image is resampled using a nearest neighbor technique to match the pixel size of the PAN image, its mean gradient was higher than that of some of the pan sharpened images.

Thus, the spatial quality of the pan sharpened images is analyzed in this work using visual analysis. The search for better quantitative spatial quality metrics is recommended for future work.

3.3.7 The True Edge Metric

Pan sharpened images are often used for automated object extraction [50] (i.e., to extract features such as roads, buildings, etc.). Nowadays, feature extraction is often automated and an edge detection operator is applied on the image to generate an edge map. A good pan sharpening technique should retain all the edges present in the PAN image in the sharpened image.

The edges in an image are calculated as follows: a Sobel edge operator [16] is applied on the image. The Sobel edge operator basically consists of two gradient operators, one in the horizontal direction and the other in the vertical direction:

www.manaraa.com

$$H_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \tag{60}$$

$$H_y = transpose(H_x), \tag{61}$$

Filtering an image u with these two gradient operators H_x and H_y results in two gradient images g_x and g_y respectively whose values at each pixel location (i,j) are given by:

$$g_x(i,j) = \sum_{m=-1}^{+1}\sum_{n=-1}^{+1} H_x(m,n)\cdot u(i+m, j+n), \tag{62}$$

$$g_y(i,j) = \sum_{m=-1}^{+1}\sum_{n=-1}^{+1} H_y(m,n)\cdot u(i+m, j+n), \tag{63}$$

g_x and g_y give the gradient of the image u in the horizontal and the vertical direction respectively. The overall gradient magnitude of the image regardless of direction is given by the Euclidean sum of the two gradient images:

$$g(i,j) = \sqrt{g_x(i,j)^2 + g_y(i,j)^2}, \tag{64}$$

The magnitude of g at each pixel location is compared against a threshold t, if it exceeds the threshold that pixel location is declared as an edge point:

$$g(i,j) \geq t, \tag{65}$$

Thus a binary edge image is created as a result that has a 1 at a pixel location if equation $g(i,j) \geq t$, (65) is satisfied and 0 otherwise. The threshold can be either chosen manually or automatically. The automatic approach was used in this study, where the threshold is set to the SNR of the gradient image.

www.manaraa.com

The above operations are performed on the PAN image and the fused images resulting from the different decomposition levels. Thus, two binary images are produced - A_{ed} (the edge image resulting from the PAN image) and F_{ed}, which is the result of applying the edge operator on the fused image. In order to evaluate the performance of the pan sharpening technique the edge images of the PAN and the pan-sharpened image are compared. The match between the edges in the PAN and the fused image is computed. Ideally, the fused image should have all the edges that were present in the PAN image. However, there will be some edges in the PAN image that will not be detected in the fused images and some edges in the fused images that will not be present in the fused images. Thus, two metrics can be computed from the edge information –

$$TE^k = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(F_{ed}^k(i,j)=1 \,\&\, A_{ed}(i,j)=1)}{\sum_{i=1}^{M}\sum_{j=1}^{N}(A_{ed}(i,j)=1)}, \tag{66}$$

$$FE^k = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(F_{ed}^k(i,j)=1 \,\&\, A_{ed}(i,j)=0)}{\sum_{i=1}^{M}\sum_{j=1}^{N}(F_{ed}^k(i,j)=1)}, \tag{67}$$

Here TE^k is the percentage of true edges found in the fused band k. A true edge in the fused image is found, if an edge is present in the fused image at a certain pixel location (i,j) and if there is also an edge at the same pixel location in the PAN image. It is assumed that all the edges present in the PAN image are true edges. Therefore, the sum of true edges found in the fused image divided by the sum of edges in the PAN image results in TE called the True Edge metric here onwards.

www.manaraa.com

FE^k is the percentage of false edges found in the fused band k called the False Edge metric. A false edge is found if an edge is present in the fused image at location (i,j), but not present in the PAN image at that location. The percentage of false edges is determined as the total number of false edges divided by the total number of edges in the fused image. However, it is possible that a true edge may not be found in the PAN image for some reason, but it is correctly found in the MS images. Thus, a false edge could be a real edge in the image. Since this metric does not convey anything about how the edge information from the PAN image is retained in the fused image, it is not discussed here.

Since this metric measures the amount of edge information in the fused image compared to the PAN image, it can be considered a spatial quality metric. During the investigation on the minimum number of decomposition levels required to merge images with a certain resolution ratio (section 5.10) it was found that for the pan sharpened images created with different number of decomposition levels, the TE values improve significantly until the spatial quality of the pan sharpened image is not sufficient as deemed by visual analysis. Thus the proposed TE metric is tested to see if it can be used to predict the minimum number of decomposition levels required to pan sharpen images with sufficient spatial quality. These concepts will be better understood upon reading sections 5.10 and 6.2.1.

www.manaraa.com

CHAPTER IV

DATA DESCRIPTION

In this chapter all the datasets that are used are described. The original and pan sharpened images for each dataset are saved in a folder that is named by the dataset itself.

4.1 File Naming Convention

Since there are many pan sharpened images for each dataset corresponding to the different methods used to pan sharpen them a systematic file naming convention must be adopted. The files containing the pan sharpened images are suffixed with the parameters of the method used to merge them. All the images are in TIFF file format. The images are named as follows:

Prefix-[Number of decomposition levels][MRA method]-[Mother wavelet | Filter]-[MS transformation option]-[Coefficient synthesis method]-[Selection rule]-[Block processing or whole dataset processing]-[Radiometric normalization option].tif. The prefix is usually a word like ‘sharp’ or ‘qb_sharp’ to convey the meaning that the file contains a pan sharpened image. The acronym for the MRA methods are ‘dwt’ for the DWT, ‘rdwt’ for the RWT, ‘awt’ for the AWT, ‘lap’ for the LP, and ‘spt’ for the SPT. The mother wavelet is identified by the string ‘db1’, ‘db2’ or in general ‘dbN’ where N is the order of the wavelet. For the AWT and the LP the Gaussian filters are identified by a

www.manaraa.com

decimal number corresponding to the value of the central coefficient of the filter (a). The binomial and QMF filters are identified by 'binomN' or 'qmfN' where N is the filter length. For the SPT this suffix tells the number of orientation bands used in the decomposition, i.e. '2ors' means two detail images were used in the pan sharpening process. The MS transformation option can be either 'sbm' or 'ihs' or 'lhs'. The 'sbm' option implies each MS band is processed separately, while 'ihs' implies that the IHS transformation was first computed and the pan sharpening applied to only this band. The coefficient synthesis method indicates whether the additive method given by 'add' or the substitutive method given by 'sub' was used. If the AABP model was used it is given by 'aabp-N-θ', the additive or substitutive version of the AABP model can be identified by the transform used ('lap' for the additive and 'rdwt' for the substitutive). The 'N' is the window size used to compute the LCC, LG and θ is the threshold. The selection rules are identified by 'srN', where N=0 for the NULL rule, 1 for the MAS rule, 2 for the WBS rule and 3 for the WBV rule and 5 for the ADD rule. If the images are processed by computing the MRA transform on the whole image this is identified by 'whl', since block processing was also implemented to handle large datasets this is identified by the string 'blk' instead. Finally the radiometric normalization option is specified by 'hm' to mean that it was done or 'nohm' which means it was omitted.

The quality metrics for each dataset are also stored in a separate text file denoted by a '-metrics.txt' trailing extension. For e.g. the pan sharpened image for the tm-erdas2 dataset - sharpbilin234-2rdwts-db2-sbm-sub-sr0-whl-nohm.tif implies that this image was pan sharpened using 2 decomposition levels of the RWT, the 'db2' wavelet was used,

www.manaraa.com

each band was pan sharpened individually and the NULL selection rule was applied. No radiometric normalization was done and the MRA transform was computed on the whole image. The quality metrics of the image are given in the file sharpbilin234-2rdwts-db2-sbm-sub-sr0-whl-nohm-metrics.txt.

The filenames of the MS and PAN images are self explanatory. They can be identified by the keywords 'ms' and 'pan' respectively and the resampled MS images contain the keyword 'res'.

4.2 Dataset – tm-erdas2

This dataset is created from a Landsat7 scene of the downtown area of Denver, Colorado, US. Since these are Landsat7 images, the spatial resolution ratio of the MS and PAN images is two. The three MS bands are band 2, 3, and 4 of the sensor. These bands correspond to the visible (bands 2, 3) and NIR portion (band 4) of the electromagnetic spectrum. The PAN image is of size 807 rows x 839 columns while the MS image is of size 404 rows x 420 columns.

4.3 Dataset – tm-vorarlberg

This dataset is also a Landsat7 scene of the Vorarlberg state of Austria. The Rhine River in the scene runs from south to north into Lake Constance. The three MS bands are band 1, 2, and 3 of the Landsat sensor. The scene is 2089 rows x 2113 columns. The original PAN and MS images are vbg_pan.tif and vbg_msb123.tif. The pan sharpened images are prefixed with 'vbg_sharp123' for images merged with the NULL rule while for merging images with the advanced rules like the MAS, WBS, etc. the MS images

www.manaraa.com

were resampled using a bilinear interpolation method and thus these images are prefixed with a 'vbg_sharpbilin123'.

4.4 Dataset – tm-vorarlberg2

This dataset is a subset of the tm-vorarlberg dataset. Since the tm-vorarlberg dataset is huge it would take up a lot of disk space to store the pan sharpened images corresponding to the different pan sharpening methods that were tried. Thus a 512x512 pixels wide region from the upper left portion of the tm-vorarlberg dataset was cropped to create this new dataset.

If the new MS image is created by taking a subset of the original low resolution MS image which is smaller in size than the PAN image then a lot of effort has to be spend in coregistering the new cropped MS and PAN images. Thus a simpler approach is used, the subset operation is performed on the resampled MS and PAN images that have the same pixel size. However, because of this the MS image of smaller size is not available to compute Wald's property 1, but this should not be a serious problem as we are the spectral fidelity is measured by comparing the resampled MS and pan sharpened images.

4.5 Dataset – qb1

This dataset is a Quickbird scene of a city in the Middle East. The first three MS bands are used here. The images have a resolution ratio of four. The PAN image size was 512x512 while the MS was 128x128 pixels.

www.manaraa.com

4.6 Dataset – qb-pyramids2

This dataset is also a Quickbird scene of an area to the north of the pyramids in Egypt. It contains urban features in the North of the imagery and mostly vegetation in the western part of the imagery. In the bottom left portion of the image there are some containers on the ground. The PAN and the resampled MS images are of size 512x512 pixels.

Since these images were also obtained by subsetting larger images like the tm-vorarlberg2 dataset the original low resolution MS image is not available for comparison.

4.7 Dataset - geovantage imagery

One of the goals of this study was to determine the minimum number of decomposition levels required to merge images with a given resolution ratio. As it was decide to use Wald's property 2 approach to aid in making the conclusions four datasets with resolution ratios 2, 3, 4, and 5 were created using Yocky's method given in section 3.3.3.5. The four datasets corresponding to resolution ratios 2, 3, 4, and 5 are called geovantage_ratio2, geovantage_ratio3, geovantage_ratio4, and geovantage_ratio5 respectively. All the datasets are of the same location - the Mississippi State University campus located in Mississippi, USA. They were created from a 1 m resolution aerial color photograph by subsetting a 600x600 pixels wide area from the original. The image contains mostly buildings, lawns, vehicles and a road running across the whole image from the top to bottom.

www.manaraa.com

CHAPTER V

RESULTS

5.1 The Effect of Radiometric Normalization

The effect of radiometric normalization on the spectral and spatial quality of the pan sharpened images was investigated by pan sharpening the datasets by first performing the histogram match of the PAN image to the MS image before doing the MRA based pan sharpening and then repeating the same process with all the parameters identical but without the histogram matching (HM) step.

5.1.1 Dataset – tm-erdas2

The quality metrics of the pan sharpened images created with and without the histogram match are given in Table 5-1. Except for the first band it is seen that the pan sharpened image created without the HM preserves the spectral fidelity better than that by the pan sharpened image with the HM. The inter band correlation is also better preserved, when the HM is avoided. However, the spatial quality of the latter is slightly better as seen in Figure 5-1. The pan sharpened images are shown with a zoom factor of two. The individual bands of both pan sharpened images were visually analyzed and it was found that the spatial quality of the individual bands differ in both the images. The

www.manaraa.com

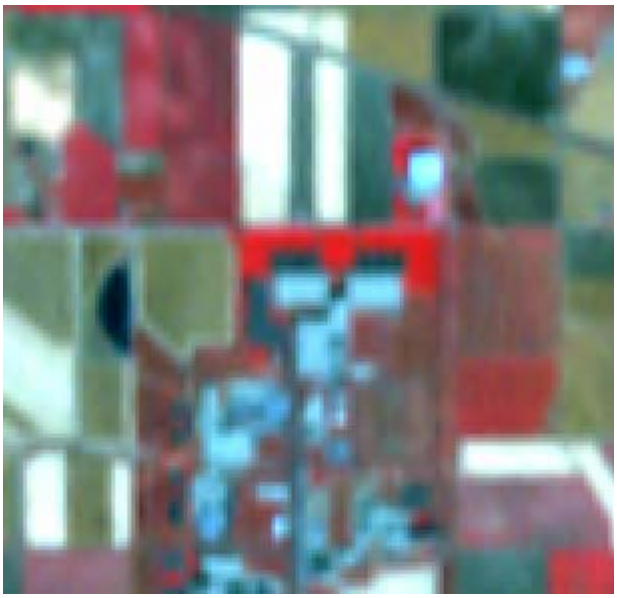
spatial quality of the first band (band 2 of TM) is better in the pan sharpened image created without the HM while the spatial quality of the other two bands (bands 3, 4 of TM) are better in the pan sharpened image created with the HM. As a result, the spatial quality of the HM pan sharpened image appears better because two of the three bands have good spatial quality. Since the goal of the pan sharpening process is to extract maximum spatial information and retain the highest spectral fidelity from the PAN image it would be desirable to pan sharpen the first band without the HM and the remaining two bands with the HM.

www.manaraa.com

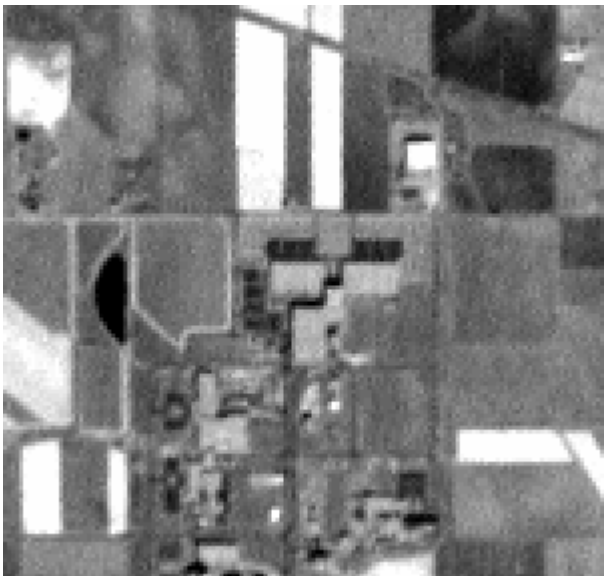
Table 5-1 Spectral quality metrics (tm-erdas2)

	HM	NO HM
$RMSE(B_1(l) \mid F_1'(l))$	1.570	1.925
$RMSE(B_2(l) \mid F_2'(l))$	2.761	2.279
$RMSE(B_3(l) \mid F_3'(l))$	2.928	2.122
$Corr(B_1(l) \mid F_1'(l))$	0.989	0.984
$Corr(B_2(l) \mid F_2'(l))$	0.991	0.994
$Corr(B_3(l) \mid F_3'(l))$	0.991	0.995
$Corr(F_1, F_2)$	0.971	0.961
$Corr(F_2, F_3)$	0.637	0.620
$Corr(F_3, F_1)$	0.643	0.630
$Corr(B_1, B_2)$	0.967	
$Corr(B_2, B_3)$	0.601	
$Corr(B_3, B_1)$	0.607	
$Corr(B_1^*, F_1)$	0.973	0.966
$Corr(B_2^*, F_2)$	0.976	0.990
$Corr(B_3^*, F_3)$	0.975	0.992
$RMSE(B_1^*, F_1)$	2.438	2.802
$RMSE(B_2^*, F_2)$	4.465	2.839
$RMSE(B_3^*, F_3)$	4.834	2.755
$Corr(B_1^*, A)$	0.829	
$Corr(B_2^*, A)$	0.822	
$Corr(B_3^*, A)$	0.897	
$Corr(F_1 \mid A)$	0.861	0.870
$Corr(F_2 \mid A)$	0.857	0.849
$Corr(F_3 \mid A)$	0.930	0.920

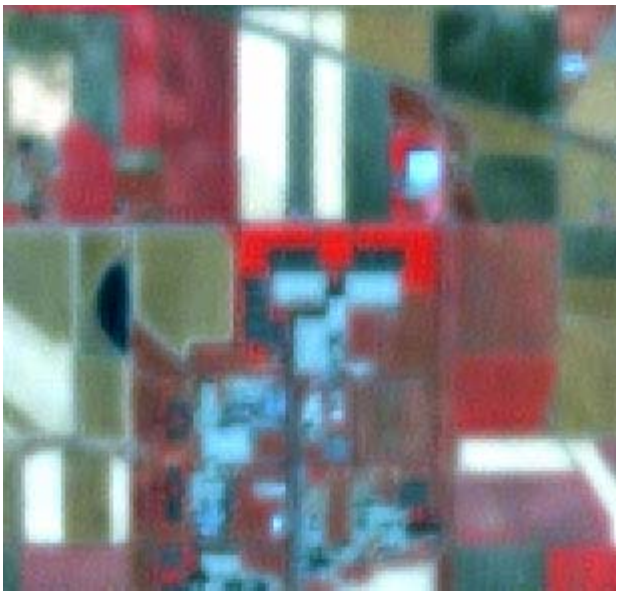
www.manaraa.com



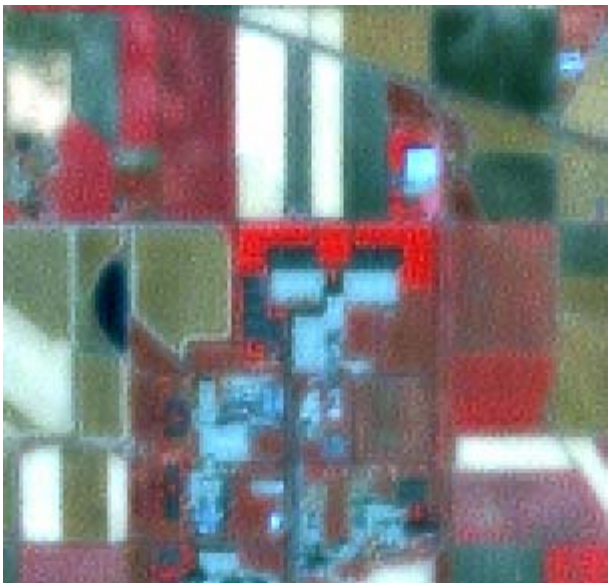
(a) The resampled MS image



(b) The PAN image



(c) Pan sharpened image without the HM



(d) Pan sharpened image with the HM

Figure 5-1 Pan sharpened images with and without the HM (tm-erdas2)

www.manaraa.com

5.1.2 Dataset – tm-vorarlberg

The pan sharpening of this data is carried out using three decomposition levels compared to only two in the above dataset. This was done because pan sharpened images produced using one or two decomposition levels were quite blurry and appear unmerged in many parts of the image. This topic is addressed later in this work in section 5.10 by making a detailed analysis of the effect of number of decomposition levels on the spectral and spatial quality of pan sharpened images.

The pan sharpened images with and without the histogram match are – vbg_sharp123-3rdwts-db2-sbm-sub-sr0-whl-hm.tif and vbg_sharp123-3rdwts-db2-sbm-sub-sr0-whl-nohm.tif. The metrics are given in Table 5-2. The spectral quality of the pan sharpened image created with the HM is slightly better. However, the spatial quality of this image is very poor. The spatial quality of each band was individually studied and it was found that for each band the quality of the pan sharpened band without the HM is better. This can be seen from Figure 5-2, where a small area of the image is shown. Many terrain features that could not be observed in the original MS image are seen in the pan sharpened image created without the histogram match, but cannot be delineated as well in the pan sharpened image created with the histogram match. Many more such features can be identified by browsing the complete imagery on the disk.

www.manaraa.com

Table 5-2 Spectral quality metrics (tm-vorarlberg)

	HM	NO HM
$RMSE(B_1(l) \mid F_1'(l))$	6.518	6.848
$RMSE(B_2(l) \mid F_2'(l))$	6.012	6.319
$RMSE(B_3(l) \mid F_3'(l))$	8.637	7.857
$Corr(B_1(l) \mid F_1'(l))$	0.943	0.937
$Corr(B_2(l) \mid F_2'(l))$	0.954	0.949
$Corr(B_3(l) \mid F_3'(l))$	0.932	0.943
$Corr(F_1, F_2)$	0.965	0.969
$Corr(F_2, F_3)$	0.981	0.982
$Corr(F_3, F_1)$	0.963	0.963
$Corr(B_1, B_2)$	0.962	
$Corr(B_2, B_3)$	0.975	
$Corr(B_3, B_1)$	0.957	
$Corr(B_1^*, F_1)$	0.930	0.924
$Corr(B_2^*, F_2)$	0.941	0.936
$Corr(B_3^*, F_3)$	0.918	0.934
$RMSE(B_1^*, F_1)$	7.334	7.553
$RMSE(B_2^*, F_2)$	6.865	7.118
$RMSE(B_3^*, F_3)$	9.546	8.476
$Corr(B_1^*, A)$	0.495	
$Corr(B_2^*, A)$	0.666	
$Corr(B_3^*, A)$	0.608	
$Corr(F_1 \mid A)$	0.566	0.591
$Corr(F_2 \mid A)$	0.730	0.745
$Corr(F_3 \mid A)$	0.694	0.701

www.manaraa.com

(a) The resampled MS image

(b) The PAN image

(c) Pan sharpened image without the HM

(d) Pan sharpened image with the HM

Figure 5-2 Pan sharpened images with and without the HM (tm-vorarlberg)

Clearly the initial histogram match step should be avoided when pan sharpening this dataset. The original PAN image (vbg_pan.tif) and the PAN image histogram matched to each band (vbg_pan_hm_b1.tif, vbg_pan_hm_b2.tif, and vbg_pan_hm_b3.tif) were analyzed. It is found that the histogram match alters/increases the DN values of the

www.manaraa.com

PAN image in such a way that the texture or details of the image are blurred or lost. This can be observed by looking at the PAN image before and after the histogram match. Since the details of the histogram matched PAN image are substituted in the pan sharpened image, it too has lost the details. One more side effect of the increased DN values of the PAN image after histogram match is that the pan sharpened image appears saturated in color in many parts of the forest. These appear as bright green spots in many parts of the image.

Thus it is seen that in the first dataset the histogram match improved the quality of the pan sharpened images while in the second case it reduced it. Though both the datasets are TM images, the first one was composed of bands 2, 3, 4 while the second one is composed of bands 1, 2, 3. The first one was taken over a mainly urban scene while the second one comprises of forests and mostly natural landscape. It is known that the PAN band does not spectrally overlap the first MS band and this may be the cause of the contradictions in the results.

5.1.3 Dataset – qb1

The pan sharpened images with and without the initial histogram match are – qb_sharp-3rdwts-db2-sbm-sub-sr0-whl-hm.tif and qb_sharp-3rdwts-db2-sbm-sub-sr0-whl-nohm.tif. Table 5-3 gives the spectral and spatial quality metrics for these two pan sharpened images.

www.manaraa.com

Table 5-3 Spectral and spatial quality metrics for the qb1 dataset

	HM	NO HM
$RMSE(B_1(l) \mid F_1'(l))$	2913	3129
$RMSE(B_2(l) \mid F_2'(l))$	2889	3078
$RMSE(B_3(l) \mid F_3'(l))$	2878	3023
$Corr(B_1(l) \mid F_1'(l))$	0.936	0.927
$Corr(B_2(l) \mid F_2'(l))$	0.941	0.934
$Corr(B_3(l) \mid F_3'(l))$	0.947	0.942
$Corr(F_1, F_2)$	0.988	0.990
$Corr(F_2, F_3)$	0.990	0.991
$Corr(F_3, F_1)$	0.962	0.965
$Corr(B_1, B_2)$	0.983	
$Corr(B_2, B_3)$	0.986	
$Corr(B_3, B_1)$	0.950	
$Corr(B_1^*, F_1)$	0.910	0.889
$Corr(B_2^*, F_2)$	0.913	0.896
$Corr(B_3^*, F_3)$	0.918	0.906
$RMSE(B_1^*, F_1)$	3599	4086
$RMSE(B_2^*, F_2)$	3633	4056
$RMSE(B_3^*, F_3)$	3711	4035
$Corr(B_1^*, A)$	0.792	
$Corr(B_2^*, A)$	0.838	
$Corr(B_3^*, A)$	0.854	
$Corr(F_1 \mid A)$	0.9	0.907
$Corr(F_2 \mid A)$	0.939	0.944
$Corr(F_3 \mid A)$	0.951	0.955

It is seen from the metrics that the histogram match resulted in a slightly better spectral fidelity. The spatial quality of both the pan sharpened images is good. Both appear sharp and have all the details of the PAN image. Thus, in this case the histogram match did not have a significant effect on the results. In preference of the higher spectral fidelity provided by the histogram match this image would be preferred over the one

www.manaraa.com

produced without the histogram match. After all, that is the main motivation for choosing the MRA based fusion technique over the conventional methods like HIS, PCA, etc.

5.1.4 Dataset – qb-pyramids2

The same experiments were conducted for this dataset by keeping the same parameters as for the above dataset. The main conclusion from this experiment indicates that the histogram match step results in slightly higher spectral fidelity. The reason for the similarity in the findings may be because both the images are of the first three bands and the PAN image's spectral range overlaps all the three bands. This indicates that the Quickbird images must be pan sharpened with the initial histogram match. The metrics for this dataset are not given in order to preserve space.

5.1.5 Summary of the Histogram Match

From the above results it was seen that for the Quickbird bands the pan sharpened images produced by applying a histogram match increased the spectral fidelity while for the TM bands it was seen that for one dataset (tm-erdas2) the histogram matching improved the results for two of the bands while for the other dataset it degraded the spatial quality in all the three bands.

For the second TM dataset (tm-vorarlberg) it is suspected that the MS and PAN images have a lot of contrast reversal. Generally, regions of contrast reversal are identified by very low correlation between the original MS and PAN images. The correlation between the MS and PAN images for the tm-erdas2, qb1, and qb-pyramids2 dataset is quite high in the range of 0.78 to 0.90 while that for tm-vorarlberg is 0.5, 0.67,

www.manaraa.com

and 0.61 in bands 1, 2, and 3 respectively. The correlation in the upper left 512x512 region of this image (the tm-vorarlberg2 dataset) was even lower – -0.175, 0.115, and 0.302. This low correlation or contrast reversal can be identified even visually. Most of the features or landscapes that appear bright in the PAN image appear dark in the MS bands and vice versa. For example, the Rhine River in the center of the image appears dark in the PAN image whereas it appears very bright in the MS bands.

Thus, from pan sharpening the above datasets it seems that the histogram matching should be performed in general which seems to increase the spectral fidelity of the pan sharpened images except when there is contrast reversal in the MS-PAN images. In a practical implementation of the pan sharpening algorithm the user should not have to specify whether histogram matching should be performed or not. Another point to be noted is that contrast reversal can occur in isolated small regions which are just a few pixels wide or in a particular feature like a river which could run across the whole image (like in the tm-vorarlberg dataset). Thus if contrast reversal is a concern the adaptive processing schemes given in [19], [34] should be implemented instead.

However, as mentioned previously these schemes could not be implemented due to scope limitations. From this point forward in the dissertation the above datasets the appropriate preprocessing is applied for each dataset. This means processing the tm-erdas2, qb1, and qb-pyramids2 datasets with a histogram match but avoiding this step for the tm-vorarlberg, and tm-vorarlberg2 datasets.

www.manaraa.com

5.2 Comparison of the Additive and Substitutive Methods

One of the main limitations of the additive merger technique is that they can only be applied to datasets whose resolution ratios are a power of two (section 2.4.4). The RWT, AWT, and LP are applied and the quality of the pan sharpened image produced by each of them is compared to that obtained from the substitutive method. For convenience the additive RWT, AWT, and LP methods will be denoted by RWT-ADD, AWT-ADD, and LP-ADD, respectively.

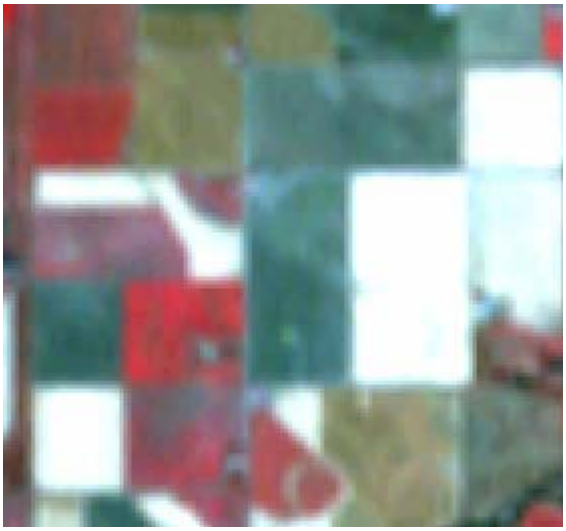
The resolution ratios of the tm-erdas2 and geovantage_ratio2 datasets are two thus one decomposition level is applied during the merging. The resolution ratios for the qb1 and geovantage_ratio4 datasets are four and thus, two decomposition levels are applied for merging them. The 'db2' wavelet is applied for the RWT-ADD method and the Gaussian filter corresponding to a=0.375 is applied for the AWT and the LP methods.

5.2.1 Dataset – tm-erdas2

The pan sharpened image created from the RWT-ADD, AWT-ADD and the LP-ADD appear satisfactorily merged at first. However, closer inspection reveals that the RWT-ADD pan sharpened image appears grainy or freckled compared to the AWT-ADD and LP-ADD pan sharpened image. This effect is shown in Figure 5-3. Figure 5-3(a) is the resampled, but unsharpened MS image while Figure 5-3(b) is the pan sharpened image produced using two decomposition levels of RWT, and the substitutive method. Figure 5-3(c) is the pan sharpened image using RWT-ADD and Figure 5-3(d) is the pan sharpened image using AWT-ADD method. The images were scaled down by a factor of two to clearly demonstrate the freckled effect in the RWT-ADD image. The LP-ADD pan

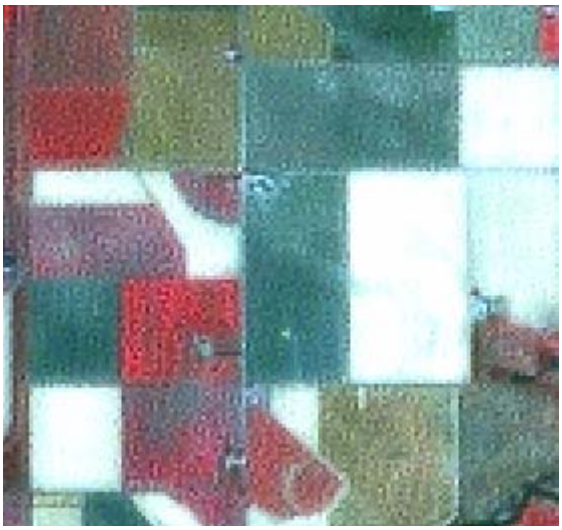
www.manaraa.com

sharpened image is not shown as it is identical to the AWT-ADD image. The spatial quality of the AWT-ADD and the LP-ADD methods is as good as that of the substitutive method.



(a) The resampled MS image

(b) Pan sharpened image with 2 RWTs and the SUB method



(c) Pan sharpened image using RWT-ADD method

(d) Pan sharpened image using AWT-ADD method

Figure 5-3 Pan sharpened image using the additive merger (tm-erdas2)

The quality metrics for the pan sharpened images created using the additive and substitutive methods are given in Table 5-4. The correlation and the RMSE between the

www.manaraa.com

MS image and the pan sharpened image cannot be computed for the LP-ADD image because in this technique there is no need to resample the MS image. Thus, only Wald's property 1 is calculated. The RMSE values for the RWT pan sharpened image are very high while that of the AWT-ADD and the LP-ADD method are small. This is proof that they have preserved the spectral information of the original MS image while enhancing its spatial quality. The AWT-ADD merger seems to best preserve the spectral fidelity. It is seen that the spectral fidelity of both LP-ADD and AWT-ADD methods is better than that of the substitutive method.

www.manaraa.com

Table 5-4 Spectral quality metrics for the additive merger (tm-erdas2)

	RWT-SUB	RWT-ADD	AWT-ADD	LP-ADD
$RMSE(B_1(l) \mid F_1'(l))$	1.57	27.94	0.920	1.392
$RMSE(B_2(l) \mid F_2'(l))$	2.761	34.86	1.717	2.661
$RMSE(B_3(l) \mid F_3'(l))$	2.928	39.65	1.826	2.908
$Corr(B_1(l) \mid F_1'(l))$	0.989	0.991	0.996	0.992
$Corr(B_2(l) \mid F_2'(l))$	0.991	0.993	0.997	0.992
$Corr(B_3(l) \mid F_3'(l))$	0.991	0.991	0.997	0.991
$Corr(F_1, F_2)$	0.971	0.967	0.969	0.972
$Corr(F_2, F_3)$	0.637	0.635	0.630	0.648
$Corr(F_3, F_1)$	0.643	0.639	0.635	0.654
$Corr(B_1, B_2)$	0.967			
$Corr(B_2, B_3)$	0.601			
$Corr(B_3, B_1)$	0.607			
$Corr(B_1^*, F_1)$	0.973	0.954	0.982	-
$Corr(B_2^*, F_2)$	0.976	0.961	0.984	-
$Corr(B_3^*, F_3)$	0.975	0.956	0.983	-
$RMSE(B_1^*, F_1)$	2.438	28.10	2.065	-
$RMSE(B_2^*, F_2)$	4.465	35.03	3.756	-
$RMSE(B_3^*, F_3)$	4.834	39.83	4.082	-
$Corr(B_1^*, A)$	0.829			
$Corr(B_2^*, A)$	0.822			
$Corr(B_3^*, A)$	0.897			
$Corr(F_1 \mid A)$	0.861	0.846	0.856	-
$Corr(F_2 \mid A)$	0.857	0.845	0.853	-
$Corr(F_3 \mid A)$	0.930	0.917	0.926	-

5.2.2 Dataset – geovantage_ratio2

The images of this dataset were also pan sharpened using the RWT-ADD, AWT-ADD, and LP-ADD methods. The parameters of the fusion process are the same as for the above dataset, except that the histogram match is not performed. The spectral quality

www.manaraa.com

metrics of the pan sharpened images are given in Table 5-5. The pan sharpened images are - sharp-1awts-0.375-sbm-add-whl-nohm.tif, sharp-1laps-0.375-sbm-add-whl-nohm.tif and sharp-1rdwts-db2-sbm-add-whl-nohm.tif (corresponding to the AWT-ADD, LP-ADD and the RWT-ADD methods, respectively). The metrics show that the AWT and LP methods result in a pan sharpened image that is closer to the original high resolution MS image. Visually all the pan sharpened images appear quite similar. However, closer inspection shows that the RWT pan sharpened image appears slightly freckled in certain portions of the image. One such portion is shown in Figure 5-4, which shows the original high resolution MS image (a), the pan sharpened image created using 3 levels of RDWT and the substitutive method (b), the pan sharpened images corresponding to the AWT-ADD (c) method, and the RWT-ADD method (d). It can be seen that the RWT-ADD image is slightly blurry near the edges of objects. This region is shown at a smaller scale in order to clearly demonstrate the poor edge information in the additively merged images clearly. The image corresponding to the additive LP method is not shown as it is similar to the AWT image (sharp-1laps-B3-sbm-add-whl-nohm.tif).

Table 5-5 Spectral quality metrics for the additive merger (geovantage_ratio2)

Wald's property 2	RWT-SUB	RWT-ADD	AWT-ADD	LP-ADD
$RMSE(B_1 \mid F_1)$	8.8	69	13	16
$RMSE(B_2 \mid F_2)$	8.5	60	13	16
$RMSE(B_3 \mid F_3)$	10	65	15	18
$Corr(B_1 \mid F_1)$	0.99	0.94	0.97	0.95
$Corr(B_2 \mid F_2)$	0.98	0.93	0.96	0.94
$Corr(B_3 \mid F_3)$	0.98	0.93	0.96	0.94

www.manaraa.com

(a) Original high resolution MS image

(b) Pan sharpened image corresponding to 3 RWT levels and substitutive method

(c) Pan sharpened image corresponding to AWT-ADD method

(d) Pan sharpened image corresponding to RWT-ADD method

Figure 5-4 Pan sharpened images from the additive method (geovantage_ratio2)

www.manaraa.com

Another interesting observation made from this dataset is that the pan sharpened images created from the additive method are not as sharp as that obtained from three decomposition levels and the substitutive method. The white border of the blue roof on the upper right portion is retained in the pan sharpened image produced from the three decomposition levels of RWT and the substitutive method, but not so clearly in the additively merged images. This region is marked in the figure with a white boundary around it. However, this can be noticed only at higher scales. Thus, it is seen that the AWT-ADD and LP-ADD images have moderate spatial quality, and it is up to the user to decide if this is acceptable for them.

5.2.3 Dataset – qb1

The three pan sharpened images corresponding to the RWT-ADD, AWT-ADD and the LP-ADD methods are - qb_sharpbilin-2rdwts-db2-sbm-add-whl-hm.tif, qb_sharpbilin-2awts-0.375-sbm-add-whl-hm.tif and qb_sharp-2laps-0.375-sbm-add-whl-hm.tif respectively. Their spectral quality metrics are given in Table 5-6. In order to do a comparison with the substitutive method, the pan sharpened image created from three decomposition levels of the RWT is used – qb_sharp-3rdwts-db2-sbm-sub-whl-hm.tif.

www.manaraa.com

Table 5-6 Spectral quality metrics for the additive merger (qb1)

	RWT-SUB	RWT-ADD	AWT-ADD	LP-ADD
$RMSE(B_1(l) \mid F_1'(l))$	2913	14848	1406	3301
$RMSE(B_2(l) \mid F_2'(l))$	2889	17066	1385	3312
$RMSE(B_3(l) \mid F_3'(l))$	2878	17226	1404	3316
$Corr(B_1(l) \mid F_1'(l))$	0.936	0.915	0.985	0.922
$Corr(B_2(l) \mid F_2'(l))$	0.941	0.919	0.987	0.926
$Corr(B_3(l) \mid F_3'(l))$	0.947	0.922	0.987	0.933
$Corr(F_1, F_2)$	0.988	0.990	0.986	0.989
$Corr(F_2, F_3)$	0.990	0.993	0.989	0.991
$Corr(F_3, F_1)$	0.962	0.972	0.959	0.964
$Corr(B_1, B_2)$	0.983			
$Corr(B_2, B_3)$	0.986			
$Corr(B_3, B_1)$	0.950			
$Corr(B_1^*, F_1)$	0.910	0.705	0.952	-
$Corr(B_2^*, F_2)$	0.913	0.708	0.953	-
$Corr(B_3^*, F_3)$	0.918	0.715	0.953	-
$RMSE(B_1^*, F_1)$	3599	14799	2672	-
$RMSE(B_2^*, F_2)$	3633	17027	2726	-
$RMSE(B_3^*, F_3)$	3711	17197	2853	-
$Corr(B_1^*, A)$	0.758			
$Corr(B_2^*, A)$	0.805			
$Corr(B_3^*, A)$	0.822			
$Corr(F_1 \mid A)$	0.9	0.803	0.852	-
$Corr(F_2 \mid A)$	0.939	0.835	0.892	-
$Corr(F_3 \mid A)$	0.951	0.850	0.909	-

Figure 5-5 shows the pan sharpened image corresponding to the RWT-ADD method. It seems to have retained the spatial quality of the PAN image very well. For example the white markers along the road appear very clearly and the crowns of the trees in the upper section of the image are also seen very clearly. However, the colors are

www.manaraa.com

excessively strong compared to the original MS image. The rooftops have a strong shade of blue while in the original MS image they are almost white or very light blue. Consequently, the spectral quality is very poor as evidenced by the metrics. Figure 5-6 shows the pan sharpened image created using the AWT-ADD technique. The spectral quality of this image is the best among the three methods used. However, the spatial quality seems insufficient as seen from some of the rooftops. The edges of the rooftops have the brown color of the adjacent ground and thus seem to be partially merged. When the original MS image is resampled to the same size as the PAN image, it causes the spectral signatures of the rooftops and adjacent ground to become mixed. Since this image is simply substituted into the approximation image of the PAN, the mixed signature is retained in the pan sharpened process. This can also be seen in the lower right portion of the image, where there appears a blur or a smeared effect around some of the automobiles. This smearing does not occur in the original PAN image.

The pan sharpened image created using the LP-ADD method is free of this effect (Figure 5-7). This is perhaps because the image does not have to be resampled prior to insertion into the pan sharpened image. The LP-ADD image appears satisfactorily merged and clearly better than the RWT-ADD and AWT-ADD images. The smearing effect is not very serious because it is difficult to detect at the original scale of the image. It only becomes apparent when the images are observed at a smaller scale. For convenience this is shown in Figure 5-8. The spectral quality of the LP-ADD image as given by Wald's property 1 is not as good as the AWT-ADD method, but it presents a good tradeoff between spectral and spatial quality.

www.manaraa.com

Figure 5-5 Pan sharpened image using RWT-ADD technique (qb1)

www.manaraa.com

Figure 5-6 Pan sharpened image using AWT-ADD technique (qb1)

www.manaraa.com

Figure 5-7 Pan sharpened image using LP-ADD technique

www.manaraa.com

(a) AWT-ADD pan sharpened image

(b) LP-ADD pan sharpened image

Figure 5-8 Spectral signatures mixing in the AWT-ADD image

5.2.4 Dataset – geovantage_ratio4

The pan sharpened images corresponding to the RWT-ADD, AWT-ADD, and the LP-ADD methods are sharp-2rdwts-db2-sbm-add-whl-nohm.tif, sharp-2awts-0.375-sbm-add-sr0-whl-nohm.tif and sharp-2laps-0.375-sbm-add-whl-nohm.tif, respectively.

As in the above dataset (section 5.2.3) the RWT-ADD image is excessively strong in color or "blotchy" in appearance. Though the AWT-ADD image initially appears sufficiently merged at first, under closer appearance, it appears blurry and unmerged around the road. The LP-ADD image is relatively better and does not appear blurry in this area. However, the road and the divider are seen in a very weak color and overall the contrast of the LP-ADD image is very poor compared to the original high resolution MS image. Table 5-7 gives the RMSE and correlation of each pan sharpened image with respect to the high resolution MS image. It shows that among the additively merged images, the AWT-ADD image most closely resembles the high resolution image followed by the LP-ADD image and finally the RWT-ADD image. However, visual

www.manaraa.com

inspection showed that the AWT-ADD pan sharpened image has some problem and it cannot be considered a good approximation of the high resolution image. A pan sharpened image using three decomposition levels and the substitutive technique was compared with these images. The spatial quality of this image was clearly better than the RWT-ADD, AWT-ADD, and the LP-ADD images. The metrics for this image are also much better than those of the additive images. Thus, for this dataset it seems that the LP-ADD image provided only a modest solution (similar to the geovantage_ratio2 dataset).

Table 5-7 Spectral quality metrics for the additive merger (geovantage_ratio4)

Wald's property 2	RWT-SUB	RWT-ADD	AWT-ADD	LP-ADD
$RMSE(B_1 \mid F_1)$	9.9	100	11	15
$RMSE(B_2 \mid F_2)$	9.6	88	11	15
$RMSE(B_3 \mid F_3)$	12	95	13	17
$Corr(B_1 \mid F_1)$	0.98	0.86	0.98	0.96
$Corr(B_2 \mid F_2)$	0.98	0.84	0.97	0.95
$Corr(B_3 \mid F_3)$	0.97	0.85	0.97	0.95

5.2.5 Summary of the Additive Method

Thus, from the above results it can be summarized that the RWT based additive merger is not suitable. This is due to the fact that it results in freckled images sometimes or too strong colors and consequently leads to a very high loss of spectral fidelity.

The AWT and LP based additive merger gave good results when pan sharpening the tm-erdas2 and geovantage_ratio2 datasets. Both the methods resulted in higher spectral fidelity compared to the substitutive method for the tm-erdas2 dataset. However, the spatial quality of the pan sharpened images created by these methods for the

www.manaraa.com

geovantage_ratio2 dataset is slightly less than that of the substitutive method. Although this difference, in the spatial quality can only be observed on closer inspection or at a higher scale.

The AWT-ADD method could not achieve good spatial quality when used for sharpening the other two datasets with a higher MS-PAN resolution ratio of four. Thus it can be said that the AWT-ADD method is only suitable for merging images with a resolution ratio of two.

The LP based additive merger is deemed the most suitable of the three additive methods studied. It did not give the blotchy appearance like the RWT and retained higher spatial quality compared to the AWT method when it was used to pan sharpen images with a resolution ratio of four. For the tm-erdas2 dataset it resulted in higher spectral fidelity than the substitutive method based on the RWT, while for the qb1 dataset the substitutive RWT method gave higher spectral fidelity. Later (through the experiments in section 5.3 that were conducted to determine the best wavelet basis for the RWT and the best filter for the AWT and LP) it was found that the 'db2' wavelet gives the highest spectral fidelity among all the Daubechies and biorthogonal wavelets that were selected for evaluation. However, for the LP it turned out that the Gaussian filter corresponding to a=0.375 is not the best one, but the 13-tap QMF. This could be the reason that the substitutive RWT method gave higher spectral fidelity for the qb1 dataset. It must be noted that the filter has an effect on only the spectral quality and not the spatial quality of the pan sharpened images.

www.manaraa.com

Thus it can be said that the LP-ADD method results in high spectral quality pan sharpened images that are comparable to those produced by the RWT based substitutive method. It was later found that the spectral fidelity of the substitutive methods can be increased by using one of the advanced selection rules given in section 2.4.5 and studied in this chapter in section 5.6. These rules cannot be applied to the additive method. Thus, the substitutive method would outperform the additive methods if the selection rules are applied.

The main conclusion from this experiment is that the additive method based on the LP is a good choice for pan sharpening, if the user can live with modest spatial quality. However, the additive method is restricted to images with a resolution ratio that is a power of two.

5.3 Effect of the Wavelet Basis

The effect of using different wavelet bases was studied using the Daubechies and biorthogonal family of wavelets. The different Daubechies wavelets used are – 'db1', 'db2', 'db5', 'db10', and 'db20'. The following biorthogonal filters were investigated: bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8.

5.3.1 Dataset – tm-erdas2

The pan sharpened images are – sharpbilin234-2rdwts-db1-sbm-sub-sr0-whl-hm.tif, sharpbilin234-2rdwts-db2-sbm-sub-sr0-whl-hm.tif, ···, and sharpbilin234-2rdwts-db20-sbm-sub-sr0-whl-hm.tif, corresponding to the 'db1', 'db2', ···, 'db20' wavelets,

www.manaraa.com

respectively. The spectral quality metrics of the pan sharpened images produced using different Daubechies wavelets are given in Table 5-8.

Table 5-8 Spectral quality metrics for different wavelets (tm-erdas2)

	'db1'	'db2'	'db5'	'db10'	'db20'
$RMSE(B_1(l) \mid F_1'(l))$	1.673	1.570	**1.557**	1.572	1.588
$RMSE(B_2(l) \mid F_2'(l))$	2.967	2.761	**2.740**	2.772	2.807
$RMSE(B_3(l) \mid F_3'(l))$	3.025	2.928	**2.948**	2.988	3.024
$Corr(B_1(l) \mid F_1'(l))$	0.988	0.989	0.989	0.989	0.989
$Corr(B_2(l) \mid F_2'(l))$	0.989	0.991	0.991	0.991	0.991
$Corr(B_3(l) \mid F_3'(l))$	0.990	0.991	0.991	0.990	0.990
$Corr(F_1, F_2)$	0.973	0.971	0.971	0.971	**0.970**
$Corr(F_2, F_3)$	0.657	0.637	0.630	0.629	**0.628**
$Corr(F_3, F_1)$	0.663	0.643	0.636	0.634	**0.634**
$Corr(B_1, B_2)$			0.967		
$Corr(B_2, B_3)$			0.601		
$Corr(B_3, B_1)$			0.607		
$Corr(B_1^*, F_1)$	0.972	0.973	0.973	0.973	0.973
$Corr(B_2^*, F_2)$	0.975	0.976	0.976	0.976	0.976
$Corr(B_3^*, F_3)$	0.975	0.975	0.975	0.974	0.974
$RMSE(B_1^*, F_1)$	2.480	**2.438**	2.447	2.463	2.477
$RMSE(B_2^*, F_2)$	4.559	**4.465**	4.480	4.510	4.538
$RMSE(B_3^*, F_3)$	4.858	**4.834**	4.870	4.904	4.931
$Corr(B_1^*, A)$			0.829		
$Corr(B_2^*, A)$			0.822		
$Corr(B_3^*, A)$			0.897		
$Corr(F_1 \mid A)$	0.869	0.861	0.858	0.857	0.856
$Corr(F_2 \mid A)$	0.866	0.857	0.854	0.853	0.852
$Corr(F_3 \mid A)$	0.934	0.930	0.928	0.927	0.927

It is seen that according to Wald's property 1 the 'db5' wavelet gives the highest spectral fidelity while according to the correlation and RMSE values between the

www.manaraa.com

resampled MS and the pan sharpened images the ‘db2’ wavelet gives the highest spectral fidelity. It is seen that the spectral fidelity of the ‘db2’ wavelet is slightly better than ‘db1’ and the other longer wavelets. Another observation is that as the wavelet length is increased after ‘db2’, the spectral fidelity decreases steadily. The inter band correlation among the pan sharpened bands is best preserved by the ‘db20’ wavelet. From the above table it is seen that the inter band correlation steadily improves as higher order wavelets are used.

Next, the biorthogonal wavelets are tested. The results of only the few wavelets which gave better results will be given here. The interested reader can study all the metrics corresponding to each wavelet (e.g. sharpbilin234-2rdwts-bior2.2-sbm-sub-sr0-whl-hm.tif, etc.) by going to each metric’s file on the disk. In summary, the biorthogonal wavelet that resulted in the highest spectral fidelity (in terms of RMSE between MS and pan sharpened image) was ‘bior2.2’. Interestingly the RMSE metrics for the ‘db2’ are also almost the same but slightly higher. It was further decided to investigate the relationship between the number of vanishing moments in the scaling function and the spectral fidelity. Since the number of vanishing moments of the analysis scaling function and the synthesis scaling function are different, this is explored by –

i) Increasing the number of vanishing moments in the analysis scaling function while keeping them constant for the synthesis scaling function. In order to understand this effect we can look at the images and metrics’ files corresponding to three sets of wavelets. The first set {‘bior1.1’, ‘bior1.3’, ‘bior1.5’} has one vanishing moment in its synthesis scaling function, the

www.manaraa.com

second set {'bior2.2', 'bior2.4', 'bior2.6', 'bior2.8'} has two vanishing moments in its synthesis scaling function, and finally {'bior3.1', 'bior3.3', 'bior3.5., 'bior3.7', 'bior3.9'} has three vanishing moments.The metrics are not tabulated here but rather the results are just stated and conclusions made. Among the three sets it was observed that, just like the Daubechies wavelets, as the vanishing moments in the analysis scaling function are increased the spectral fidelity improved and then dropped again. If the RMSE between the original and the pan sharpened image is used as the metric for spectral fidelity the wavelets that gave the best spectral fidelity in each set were – 'bior1.3', 'bior2.2', and 'bior3.3'. The inter band correlation was found to improve for the higher order wavelets.

ii) Increasing the number of vanishing moments in the synthesis scaling function while keeping them constant in the analysis scaling function. The wavelets to be tested were divided into five sets – {'bior1.1', 'bior3.1'}, {'bior1.3', 'bior3.3'}, {'bior1.5', 'bior3.5'}, {'bior2.4', 'bior4.4'}, and {'bior2.8', 'bior6.8'}. In the first three sets the wavelets with higher number of vanishing moments in the synthesis scaling function gave better results while for the latter two the wavelets with lower num number of vanishing moments in the synthesis scaling function gave better results. Thus, a definite statement cannot be made. One reason for this anomaly may be that a set of two is too small to make any kind of inference. For example, in the Daubechies wavelet set if we had looked at only 'db1' and 'db2' one would think that higher order

www.manaraa.com

wavelets will give better spectral fidelity while in reality the spectral fidelity deteriorated for the longer wavelets. However, consistent with the previous observations the inter band correlation improved with the increase in the wavelet order.

5.3.2 Dataset – tm-vorarlberg

The spectral quality metrics of the pan sharpened images produced using different wavelets are given in Table 5-9.

www.manaraa.com

Table 5-9 Spectral quality metrics for different wavelets (tm-vorarlberg)

	'db1'	'db2'	'db5'	'db10'	'db20'
$RMSE(B_1(l) \mid F_1'(l))$	7.212	6.848	**6.801**	6.851	6.906
$RMSE(B_2(l) \mid F_2'(l))$	6.596	6.319	**6.293**	6.338	6.384
$RMSE(B_3(l) \mid F_3'(l))$	8.275	7.857	**7.799**	7.856	7.918
$Corr(B_1(l) \mid F_1'(l))$	0.930	0.937	0.938	0.937	0.936
$Corr(B_2(l) \mid F_2'(l))$	0.944	0.949	0.950	0.949	0.948
$Corr(B_3(l) \mid F_3'(l))$	0.936	0.943	0.944	0.943	0.942
$Corr(F_1, F_2)$	0.971	0.969	0.969	**0.968**	**0.968**
$Corr(F_2, F_3)$	0.983	0.982	0.981	**0.981**	**0.981**
$Corr(F_3, F_1)$	0.964	0.963	0.962	**0.962**	**0.962**
$Corr(B_1, B_2)$			0.962		
$Corr(B_2, B_3)$			0.975		
$Corr(B_3, B_1)$			0.957		
$Corr(B_1^*, F_1)$	0.917	0.924	0.925	0.925	0.924
$Corr(B_2^*, F_2)$	0.931	0.936	0.937	0.937	0.936
$Corr(B_3^*, F_3)$	0.927	0.934	0.935	0.934	0.933
$RMSE(B_1^*, F_1)$	7.882	7.553	**7.513**	7.559	7.610
$RMSE(B_2^*, F_2)$	7.366	7.118	**7.096**	7.136	7.179
$RMSE(B_3^*, F_3)$	8.863	8.476	**8.426**	8.480	8.539
$Corr(B_1^*, A)$			0.495		
$Corr(B_2^*, A)$			0.666		
$Corr(B_3^*, A)$			0.608		
$Corr(F_1 \mid A)$	0.612	0.591	0.581	0.578	0.576
$Corr(F_2 \mid A)$	0.761	0.745	0.738	0.736	0.734
$Corr(F_3 \mid A)$	0.721	0.701	0.691	0.687	0.686

From the metrics it can be seen that the 'db5' wavelet best observes Wald's property 1 and the inter band correlation is better preserved by longer wavelets in general. The spectral fidelity according to the RMSE between the resampled and sharpened MS images is also better preserved by the 'db5' wavelet.

www.manaraa.com

The analysis of the metrics of the pan sharpened images created by using the biorthogonal wavelets is similar to the above dataset. The biorthogonal wavelets do not give a higher spectral fidelity than the 'db2' wavelet. Increasing the number of vanishing moments in the analysis scaling function improves the spectral fidelity at first but then it remains about the same or in some cases deteriorated slightly. Since there does not seem to be any advantage of the biorthogonal wavelets their analysis is discontinued here onwards.

5.3.3 Dataset – qb1

The spectral quality metrics of the pan sharpened images produced using different wavelets are given in Table 5-10.

www.manaraa.com

Table 5-10 Spectral quality metrics (qb1)

	'db1'	'db2'	'db5'	'db10'	'db20'
$RMSE(B_1(l) \mid F_1'(l))$	3076.698	2913.435	2850.800	**2844.484**	2848.002
$RMSE(B_2(l) \mid F_2'(l))$	3036.903	2889.491	2829.960	**2822.475**	2824.342
$RMSE(B_3(l) \mid F_3'(l))$	3019.591	2877.918	2823.664	**2817.758**	2820.270
$Corr(B_1(l) \mid F_1'(l))$	0.929	0.936	0.939	0.940	0.939
$Corr(B_2(l) \mid F_2'(l))$	0.935	0.941	0.943	0.944	0.944
$Corr(B_3(l) \mid F_3'(l))$	0.941	0.947	0.949	0.949	0.949
$Corr(F_1, F_2)$	0.989	0.988	0.988	**0.987**	**0.987**
$Corr(F_2, F_3)$	0.991	0.990	0.989	**0.989**	**0.989**
$Corr(F_3, F_1)$	0.965	0.962	0.960	**0.960**	**0.960**
$Corr(B_1, B_2)$			0.983		
$Corr(B_2, B_3)$			0.986		
$Corr(B_3, B_1)$			0.951		
$Corr(B_1^*, F_1)$	0.906	0.910	0.910	0.909	0.909
$Corr(B_2^*, F_2)$	0.911	0.913	0.913	0.913	0.912
$Corr(B_3^*, F_3)$	0.916	0.918	0.917	0.917	0.917
$RMSE(B_1^*, F_1)$	3646.545	**3598.991**	3604.041	3617.774	3630.398
$RMSE(B_2^*, F_2)$	3661.803	**3632.511**	3641.138	3653.484	3664.247
$RMSE(B_3^*, F_3)$	3736.388	**3711.201**	3722.669	3735.364	3746.199
$Corr(B_1^*, A)$			0.792		
$Corr(B_2^*, A)$			0.838		
$Corr(B_3^*, A)$			0.854		
$Corr(F_1 \mid A)$	0.910	0.900	0.894	0.892	0.891
$Corr(F_2 \mid A)$	0.946	0.939	0.934	0.933	0.932
$Corr(F_3 \mid A)$	0.958	0.951	0.947	0.946	0.945

For this dataset the Wald's property 1 is best preserved by the 'db10' wavelet and the inter band correlation best preserved by longer wavelets than the shorter wavelets. Once again the 'db2' wavelet results in the minimum RMSE between the resampled and the sharpened MS images and it decreases as the longer wavelets are applied.

www.manaraa.com

5.3.4 Dataset – qb-pyramids2

The spectral quality metrics of the pan sharpened images produced using different wavelets are given in Table 5-11.

Table 5-11 Spectral quality metrics (qb-pyramids2)

	'db1'	'db2'	'db5'	'db10'	'db20'
$Corr(F_1, F_2)$	0.992	0.991	**0.990**	**0.990**	**0.990**
$Corr(F_2, F_3)$	0.993	0.993	**0.992**	**0.992**	**0.992**
$Corr(F_3, F_1)$	0.980	0.977	**0.976**	**0.976**	**0.976**
$Corr(B_1, B_2)$	0.986				
$Corr(B_2, B_3)$	0.989				
$Corr(B_3, B_1)$	0.968				
$Corr(B_1^*, F_1)$	0.923	0.925	0.925	0.924	0.924
$Corr(B_2^*, F_2)$	0.928	0.929	0.929	0.929	0.928
$Corr(B_3^*, F_3)$	0.930	0.932	0.931	0.931	0.931
$RMSE(B_1^*, F_1)$	3535.235	**3486.659**	3490.034	3503.715	3515.747
$RMSE(B_2^*, F_2)$	3450.662	**3417.111**	3424.223	3437.495	3448.492
$RMSE(B_3^*, F_3)$	3642.574	**3607.764**	3616.609	3631.526	3643.918
$Corr(B_1^*, A)$	0.776				
$Corr(B_2^*, A)$	0.822				
$Corr(B_3^*, A)$	0.829				
$Corr(F_1 \mid A)$	0.880	0.871	0.866	0.865	0.864
$Corr(F_2 \mid A)$	0.917	0.910	0.906	0.905	0.904
$Corr(F_3 \mid A)$	0.920	0.914	0.910	0.909	0.908

For this dataset the Wald's property 1 could not be calculated because of the way in which the dataset was created (for details see section 4.6). However, since this property is studied for the remaining three datasets it is sufficient as the behavior of the wavelets with respect to this property is understood in general. The inter band correlation is better preserved by longer wavelets than the shorter wavelets. As observed in the

www.manaraa.com

previous datasets the spectral fidelity is best preserved by the ‘db2’ wavelet and it decreases as the longer wavelets are applied.

5.3.5 Summary of the Effect of Wavelet Filters

From the above experiments it can be seen that the spectral fidelity is better preserved in most cases by ‘db2’ or ‘db5’ in one of the cases. Since the intermediate wavelets (‘db3’, ‘db4’) were not used it cannot be said with complete confidence that ‘db2’ was the best among the Daubechies family of wavelets. However, a statement can be made in general about the performance of the wavelet filters as a function of the number of vanishing moments - wavelets with fewer vanishing moments give better spectral fidelity. Though Wald’s property 1 and the RMSE and CC between the resampled MS and pan sharpened images are intended for the same purpose – to measure the spectral similarity of the pan sharpened image with respect to the original MS image, they give slightly different answers. Wald’s property 1 indicates that slightly longer wavelets (‘db5’ or ‘db10’) are better for preserving the spectral fidelity, while the CC and RMSE between the original MS and pan sharpened images indicate that the shorter wavelets (‘db2’) are preferable. Nevertheless the difference in the conclusions that can be made based on the two different sets of metrics is minimal. Both metrics indicate that the spectral fidelity first improves slightly as the wavelet filter length is increased, but starts deteriorating after a certain length.

The correlation between the original MS and PAN images and the correlation between these bands after pan sharpening is performed is given in the bottom two rows in each table. The correlation between the pan sharpened and PAN images is highest for the

www.manaraa.com

'db1' wavelet and least for the 'db20' wavelet, thus it can be concluded that the shorter wavelets integrate a higher amount of information from both the PAN and the MS images.

5.4 Effect of Different Kernels on the LP Based Fusion

Various categories of filters were used to pan sharpen the datasets using the LP. In order to observe only the effect of the filter the NULL selection rule and single band processing were applied. The filters compared were binomial filters, Quadrature Mirror Filters (QMF) and the length 5 Gaussian kernels as given by Burt et al. [Burt, 1983]. In this section since many filters are compared the detailed results for each dataset will not be tabulated for conciseness, the interested reader can look at the metrics files on the disk.

5.4.1 Binomial Filters

Binomial filters of order one to twelve were used to do the pan sharpening. From experiments in the datasets tm-erdas2, tm-vorarlberg2, qb1, and qb-pyramids2 it is seen that the shorter binomial filters provide better spectral fidelity while the spatial quality of all pan sharpened images seemed sufficient and identical. One exception noted was that the first order or 2-tap binomial filter gave artifacts in the pan sharpened image. The images appeared pixilated and had the stair step effect almost all over the image. However, one exception is the tm-erdas2 dataset where this effect was hardly visible. It was found that as the filter order was increased the spectral fidelity decreased. The 3-tap (2^{nd} order) binomial filter resulted in the highest spectral fidelity among all the binomial

www.manaraa.com

filters. Another interesting observation is that the even length filters have very poor spectral fidelity. For e.g. the spectral fidelity of the 4-tap binomial filter was much poor compared to the 5-tap or even the 7-tap filter. This maybe because of the asymmetry of even length filters.

5.4.2 Quadrature Mirror Filters

Application of the QMF filters shows that the odd length QMF filters give good results but the even length QMF filters give artifacts. These are not very severe but they do exist in some regions of the image. These do not appear pixilated or stair stepped but rather seem to be occurring due to the smearing of spectral signatures of sharp objects and thus result in poor spectral quality. The QMF filters used were 5, 8, 9, 12, 13 and 16 tap respectively. It is also seen that as the length of the odd length QMF filters is increased the spectral quality improves. Comparison between the best binomial filter ('binom3') and the best QMF filter ('qmf13') showed that the QMF filter gave higher spectral fidelity. Even the smallest 'qmf5' filter gave better spectral quality than the best 'binom3' filter. Thus it can be concluded that odd length QMF filters must be preferred over the binomial filters.

5.4.3 Gaussian Filters

The length-5 Gaussian kernel proposed by Burt et al. [5] was used. Five different Gaussian kernels were obtained by varying the weights of the filter coefficients subject to certain constraints. The different filter coefficients obtained by varying the weights of the coefficient are given in Table 5-11. In the table each column corresponds to the filter

www.manaraa.com

coefficients obtained by changing the weight central coefficient – a, from 0.3 to 0.6. When 'a' is increased beyond 0.5 it is seen that the border coefficients become negative and the filter becomes trimodal.

Table 5-12 Various Gaussian filters

Filter Coefficients	a=0.3	a=0.375	a=0.4	a=0.5	a=0.6
w(-2)	0.10	0.0625	0.05	0	-0.05
w(-1)	0.25	0.25	0.25	0.25	0.25
w(0)	0.30	0.375	0.40	0.5	0.6
w(1)	0.25	0.25	0.25	0.25	0.25
w(2)	0.10	0.0625	0.05	0	-0.05

The effect of varying a from 0.3 to 0.6 is that the spectral fidelity increases as 'a' increases. However, it was seen that the pan sharpened images corresponding to a=0.6 filter have artifacts. These artifacts are quite minor and only a closer inspection identifies them. The only exception was the tm-erdas2 dataset otherwise these artifacts were seen in all the other datasets. These artifacts seem to be caused by the trimodal nature of the Gaussian function when a=0.6. Similarly artifacts were observed when this filter (Gaussian, a=0.6) was used to fuse two out of focus images. Thus the Gaussian filter corresponding to a=0.6 will not be considered further for LP based fusion. Thus the filter corresponding to a=0.5 seems the most suitable among the Gaussian filters.

The Gaussian filter was compared to the other binomial and QMF filters. Only the best filters from each set were chosen i.e. the 5-tap Gaussian corresponding to a=0.5, the 3-tap binomial filter and the 13-tap QMF filter. Even the 5-tap QMF filter is better than the Gaussian filters, while the quality metrics of the 3-tap binomial filter are almost equal to that of the Gaussian filter.

www.manaraa.com

5.5 Effect of Different Kernels on the AWT Based Fusion

The same filters as those used for the LP based fusion were evaluated for the AWT based pan sharpening. The comparisons were done using the NULL rule.

5.5.1 Binomial Filters

Most of the observations made for this filter family are similar to those obtained by applying the LP. One exception noted is that the 2-tap binomial filter does not give artifacts, this is perhaps a result of avoiding the decimation step. The even length filters resulted in very poor spectral fidelity compared to the odd length filters. Under this transform the 5-tap binomial filter resulted in the best spectral fidelity in the Quickbird datasets qb1, and qb-pyramids2 while the 3-tap binomial filter gave the highest spectral fidelity for the tm-erdas2, and tm-vorarlberg2 datasets.

5.5.2 Quadrature Mirror Filters

It was found that in general the odd length QMF filters gave much better results than the even length filters. Among the odd length filters the longer filters gave higher spectral fidelity. The 13-tap QMF gave the highest spectral fidelity compared to other QMFs, except for the tm-vorarlberg2 dataset where the 9-tap QMF gave higher spectral fidelity.

5.5.3 Gaussian Filters

A study of the Gaussian filters showed that as ‘a’ was increased from 0.3 to 0,5 the spectral fidelity of the pan sharpened images improved but it deteriorated for a=0.6.

www.manaraa.com

The spectral fidelity of the images produced by the a=0.6 Gaussian filter is worse than that produced by the filter corresponding to a=0.3. However, no artifacts are seen in the pan sharpened image produced by the a=0.6 Gaussian filter. One exception to this is the tm-erdas2 dataset in which the spectral fidelity of the pan sharpened image corresponding to the a=0.6 filter was higher than the a=0.5 filter. In the qb-pyramids2 dataset it was seen that the color of some of the buildings and trees had spilled over to the neighboring regions in the pan sharpened image corresponding to this filter. This can be seen by looking at the file sharp2-3awts-0.6-sbm-sub-sr0-whl-hm.tif.

5.5.4 Summary

Comparing the behavior of the filters applied under the AWT with their behavior under the LP it is seen that the behavior showed consistency under the LP. For e.g. the 3-tap binomial filter was superior to the rest of the binomial filters under the LP while under the AWT sometimes the 5-tap was better than the rest for two of the datasets.

An inter comparison of the spectral fidelity metrics of the different filter families for the LP showed that the best QMF (13-tap) outperformed the best binomial filter (3-tap) and the 3-tap binomial filter outperformed the best Gaussian filter (a=0.5) for all the datasets studied. Under the AWT no such consistent pattern was observed. However, the a=0.5 Gaussian filter (a=0.6 in case of tm-erdas2) gave higher spectral fidelity compared to the best QMF or binomial filter always.

www.manaraa.com

5.6 The Selection Rules

Various selection rules in the context of the substitutive method were described in section 2.4.5 that could be used to increase the spectral fidelity of the pan sharpened images, out of which three primary rules are compared here – MAS, WBS, and WBV with NULL rule as the reference. The RWT will be used for the pan sharpening although the results should be applicable for the other transforms too.

In order to apply the MAS, WBS and WBV rules the MS images must be resampled to the PAN image's size using a resampling technique that results in a smooth image. A bilinear or bicubic interpolation technique is suitable for this purpose but not the nearest neighbor technique since it results in a blocky and very pixilated looking resampled image. If the NN resampled MS image is used, the output pan sharpened image is also blocky and pixilated in certain areas where the coefficients of the MS image are chosen over the coefficients of the PAN image by the selection rule. The WBV or WBS schemes can be applied on a sliding window of 3x3, 5x5 or in general *mxm* pixels. It was found for the WBS and WBV rules that as the window size is increased the spectral fidelity starts to decrease, thus the optimal window size is found to be 3x3.

First the selection rules will be evaluated and the best rule will be compared to the advanced model methods i.e. the AABP model.

5.6.1 Dataset – tm-erdas2

Table 5-13 shows the spectral quality metrics of the pan sharpened images of this dataset obtained by applying the four selection rules. The four pan sharpened images corresponding to the NULL, MAS, WBS and WBV rules are - sharpbilin234-2rdwts-db2-

www.manaraa.com

sbm-sub-sr0-whl-hm.tif, sharpbilin234-2rdwts-db2-sbm-sub-sr1-whl-hm.tif, sharpbilin234-2rdwts-db2-sbm-sub-sr2-w3-whl-hm.tif and sharpbilin234-2rdwts-db2-sbm-sub-sr3-w3-whl-hm.tif respectively.

Table 5-13 Spectral quality metrics for various rules (tm-erdas2)

	NULL	MAS	WBS	WBV
$RMSE(B_1(l) \mid F_1'(l))$	1.570	**1.306**	1.383	1.428
$RMSE(B_2(l) \mid F_2'(l))$	2.761	**2.414**	2.53	2.610
$RMSE(B_3(l) \mid F_3'(l))$	2.928	**2.715**	2.772	2.842
$Corr(B_1(l) \mid F_1'(l))$	0.989	**0.992**	0.991	0.991
$Corr(B_2(l) \mid F_2'(l))$	0.991	**0.993**	0.992	0.992
$Corr(B_3(l) \mid F_3'(l))$	0.991	**0.992**	0.992	0.991
$Corr(F_1, F_2)$	0.971	**0.970**	0.970	0.970
$Corr(F_2, F_3)$	0.637	**0.632**	0.632	0.634
$Corr(F_3, F_1)$	0.643	**0.637**	0.638	0.639
$Corr(B_1, B_2)$	0.967			
$Corr(B_2, B_3)$	0.601			
$Corr(B_3, B_1)$	0.607			
$Corr(B_1^*, F_1)$	0.973	**0.978**	0.977	0.976
$Corr(B_2^*, F_2)$	0.976	**0.979**	0.979	0.977
$Corr(B_3^*, F_3)$	0.975	**0.978**	0.977	0.976
$RMSE(B_1^*, F_1)$	2.438	**2.229**	2.273	2.341
$RMSE(B_2^*, F_2)$	4.465	**4.183**	4.262	4.370
$RMSE(B_3^*, F_3)$	4.834	**4.613**	4.670	4.769
$Corr(B_1^*, A)$	0.829			
$Corr(B_2^*, A)$	0.822			
$Corr(B_3^*, A)$	0.897			
$Corr(F_1 \mid A)$	0.861	0.858	0.858	0.859
$Corr(F_2 \mid A)$	0.857	0.855	0.855	0.856
$Corr(F_3 \mid A)$	0.930	0.929	0.928	0.929

www.manaraa.com

5.6.2 Dataset – tm-vorarlberg

Table 5-14 gives the spectral quality metrics for this dataset. The four pan sharpened images corresponding to the NULL, MAS, WBS and WBV rules are - vbg_sharp123-3rdwts-db2-sbm-sub-sr0-whl-nohm.tif, vbg_sharpbilin123-3rdwts-db2-sbm-sub-sr1-whl-nohm.tif, vbg_sharpbilin123-3rdwts-db2-sbm-sub-sr2-w3-whl-nohm.tif and vbg_sharpbilin123-3rdwts-db2-sbm-sub-sr3-w3-whl-nohm.tif respectively.

www.manaraa.com

Table 5-14 Spectral quality metrics for various rules (tm-vorarlberg)

	NULL	MAS	WBS	WBV
$RMSE(B_1(l) \mid F_1'(l))$	6.848	**5.40**	5.529	5.563
$RMSE(B_2(l) \mid F_2'(l))$	6.319	**4.968**	5.124	5.185
$RMSE(B_3(l) \mid F_3'(l))$	7.857	**5.265**	5.394	5.413
$Corr(B_1(l) \mid F_1'(l))$	0.937	**0.962**	0.960	0.959
$Corr(B_2(l) \mid F_2'(l))$	0.949	**0.969**	0.967	0.966
$Corr(B_3(l) \mid F_3'(l))$	0.943	**0.975**	0.973	0.973
$Corr(F_1, F_2)$	0.969	**0.967**	0.967	0.967
$Corr(F_2, F_3)$	0.982	**0.975**	0.974	0.974
$Corr(F_3, F_1)$	0.963	**0.958**	0.957	0.957
$Corr(B_1, B_2)$	0.963			
$Corr(B_2, B_3)$	0.977			
$Corr(B_3, B_1)$	0.958			
$Corr(B_1^*, F_1)$	0.924	**0.953**	0.952	0.952
$Corr(B_2^*, F_2)$	0.936	**0.961**	0.960	0.959
$Corr(B_3^*, F_3)$	0.934	**0.970**	0.970	0.970
$RMSE(B_1^*, F_1)$	7.553	**5.990**	6.051	6.084
$RMSE(B_2^*, F_2)$	7.118	**5.632**	5.718	5.777
$RMSE(B_3^*, F_3)$	8.476	**5.720**	5.748	5.767
$Corr(B_1^*, A)$	0.495			
$Corr(B_2^*, A)$	0.666			
$Corr(B_3^*, A)$	0.608			
$Corr(F_1 \mid A)$	0.591	0.578	0.578	0.577
$Corr(F_2 \mid A)$	0.745	0.734	0.734	0.734
$Corr(F_3 \mid A)$	0.701	0.677	0.676	0.676

5.6.3 Dataset – qb1

Table 5-15 gives the spectral quality metrics for this dataset. The four pan sharpened images corresponding to the NULL, MAS, WBS and WBV rules are – qb_sharp-3rdwts-db2-sbm-sub-sr0-whl-hm.tif, qb_sharpbilin-3rdwts-db2-sbm-sub-sr1-

www.manaraa.com

whl-hm.tif, qb_sharpbilin-3rdwts-db2-sbm-sub-sr2-w3-whl-hm.tif and qb_sharpbilin-3rdwts-db2-sbm-sub-sr3-w3-whl-hm.tif respectively.

Table 5-15 Spectral quality metrics for various rules (qb1)

	NULL	MAS	WBS	WBV
$RMSE(B_1(l) \mid F_1'(l))$ $RMSE(B_2(l) \mid F_2'(l))$ $RMSE(B_3(l) \mid F_3'(l))$	2913.435 2889.491 2877.918	**1855.394** **1820.597** **1845.612**	1923.831 1898.360 1921.780	2005.936 1986.232 2012.238
$Corr(B_1(l) \mid F_1'(l))$ $Corr(B_2(l) \mid F_2'(l))$ $Corr(B_3(l) \mid F_3'(l))$	0.936 0.941 0.947	**0.975** **0.977** **0.978**	0.973 0.975 0.977	0.970 0.973 0.974
$Corr(F_1, F_2)$ $Corr(F_2, F_3)$ $Corr(F_3, F_1)$	0.988 0.990 0.962	**0.987** **0.989** **0.959**	0.987 0.989 0.958	0.987 0.989 0.959
$Corr(B_1, B_2)$ $Corr(B_2, B_3)$ $Corr(B_3, B_1)$	0.983 0.986 0.950			
$Corr(B_1^*, F_1)$ $Corr(B_2^*, F_2)$ $Corr(B_3^*, F_3)$	0.910 0.913 0.918	**0.941** **0.944** **0.944**	0.940 0.942 0.943	0.937 0.939 0.940
$RMSE(B_1^*, F_1)$ $RMSE(B_2^*, F_2)$ $RMSE(B_3^*, F_3)$	**3598.991** **3632.511** **3711.201**	**2854.547** **2867.469** **3003.181**	2876.325 2892.723 3027.5	2941.363 2964.019 3101.738
$Corr(B_1^*, A)$ $Corr(B_2^*, A)$ $Corr(B_3^*, A)$	0.792 0.838 0.854			
$Corr(F_1 \mid A)$ $Corr(F_2 \mid A)$ $Corr(F_3 \mid A)$	0.900 0.939 0.951	0.859 0.9 0.915	0.860 0.901 0.916	0.862 0.903 0.918

www.manaraa.com

5.6.4 Performance Evaluation of the Selection Rules

The MAS, WBS and WBV rules show significant improvement in the spectral fidelity compared to the NULL rule. The MAS rule results in the highest spectral fidelity followed by the WBS and the WBV rule. The performance of the WBS rule is better than the WBV rule more close to the MAS rule. This maybe because this rule uses the selection action like the MAS rule when the match measure is below the threshold and since the threshold was set quite high at 0.75 it must be resulting in selection most of the time. The only difference is that the salience or energy of a coefficient is measured over a small neighborhood around the pixel rather than only the individual coefficient as in the MAS scheme. When the match measure exceeds the threshold the weighted average of the two coefficients is taken thus further retaining the information coming from the MS image.

A relative comparison between the metrics of the MAS and the NULL rule was done for the three datasets. The same was done for the comparison of MAS and WBV rules. The relative improvement in the spectral fidelity due to the MAS rule is given in Table 5-16. The three values in each cell are for the three bands. It is seen that the use of MAS over NULL rule shows significant improvement particularly in the last two datasets while it is smaller in the first dataset. The improvement from MAS rule over the WBV rule is very less for the last two datasets, however for the first dataset the relative improvement due to MAS over NULL and WBV is about the same magnitude. In any case the MAS rule should be preferred due to its lower computational complexity and better spectral fidelity.

www.manaraa.com

Table 5-16 Relative improvement of MAS over NULL and WBS rules

Dataset	MAS over NULL	MAS over WBV
tm-erdas2	8.57 6.32 4.57	4.78 4.28 3.27
tm-vorarlberg	20.69 20.88 32.52	1.55 2.51 0.81
qb1	20.68 21.06 19.08	2.95 3.26 3.18

The spatial quality of the pan sharpened images produced by using all the rules seems similar and thus the MAS rule must be favored since it always results in the highest spectral fidelity and is computationally fast. However on close inspection of some of the pan sharpened images it was found that the MAS, WBS and WBV rules have some problems. In the tm-vorarlberg dataset it was found that using the MAS, WBS and WBV scheme resulted in pan sharpened images that had pixilated edges, i.e. it appeared unsatisfactorily merged in some regions. In an attempt to eliminate this effect in the WBV scheme the window size was increased to 5x5 without any success. This phenomenon is illustrated in Figure 5-9, which shows the region around the Rhine River. The interested reader can look at all the three images corresponding to the MAS, WBS, and WBV rules and see that this effect is observed along the entire length of the river from North to South. The pan sharpened image corresponding to the WBS scheme is not shown as it has similar shortcomings like the MAS and WBV rules.

In the other qb1 dataset some of the automobiles on the road in the bottom right hand corner have a smeared effect which is not present in the original PAN image. The

www.manaraa.com

pan sharpened image created with the NULL rule is free of this effect. Although this problem is not very serious because it can hardly be observed at the original scale of the image but quite obvious when the images are observed at a smaller scale or only to a keen observer, this is shown in Figure 5-10. The pan sharpened image corresponding to the WBS rule is not shown here as it is similar to the MAS or WBV images. The original MS images when resampled have a slightly smeared effect around sharp structures that have small spectral signatures. This smearing seems to be retain in the pan sharpened images created using the MAS, WBS or WBV rules because its energy exceeds the energy of the coefficients coming from the PAN image.

Though no such effect was observed for the tm-erdas2 dataset it shows that the selection rules should be applied with caution as they can give slight artifacts in the pan sharpened images sometimes. This also goes on to demonstrate the importance of working with more than one dataset.

www.manaraa.com

(a) Pan sharpened image (NULL rule)

(b) The PAN image

(c) Pan sharpened image (MAS rule)

(d) Pan sharpened image (WBV rule)

Figure 5-9 Pan sharpened images from using various selection rules (tm-vorarlberg)

www.manaraa.com

(a) The PAN image

(b) The pan sharpened image (NULL rule)

(c) Pan sharpened image (MAS rule)

(d) Pan sharpened image (WBV rule)

Figure 5-10 Pan sharpened images from using various selection rules (qb1)

The above datasets were also pan sharpened using the AWT scheme and the NULL, MAS, WBS and WBV rules. The findings are similar to those obtained from the RWT method. The spectral fidelity of the MAS rule was highest followed by the WBS and WBV rules. The problems with the smearing of small spectral signatures in the qb1 dataset and the stair stepped effect in the tm-vorarlberg data were also seen, which shows that changing the MRA transform does not improve the spatial quality.

www.manaraa.com

5.7 The AABP Model

The AABP model was described in section 2.4.6.2. There are two variants of this model, AABP-ADD based on the LP which is additive and the other AABP-SUB based on the RWT which is substitutive. Both the models are implemented and studied here.

5.7.1 AABP-ADD

In the additive AABP model only the PAN image is decomposed *n* times where *n* is the resolution ratio of the MS and PAN images. However, initial results with this scheme were discouraging. The pan sharpened images had artifacts all over the images.

In an attempt to eliminate the artifacts I tried some variations of the scheme without any success. For example, different thresholds were used – θ=-1, θ=0.2, θ=0.4, and θ=0.8 to control the injection of the details in to the pan sharpened image. A threshold of -1 corresponds to unconditional injection, while as the threshold increases fewer coefficients are added to the MS image. Another variation that was tried was to omit the scaling of the detail coefficients by the LG factor but this too did not help to improve the results. The interested reader can look at the pan sharpened images in the tm-erdas2 and the geovantage_ratio2 datasets corresponding to this method (e.g., sharp-1laps-elp11-sbm-aabp-9--1-whl-nohm.tif, sharp-1laps-elp11-sbm-aabp-9-0.8-whl-nohm.tif in the tm-erdas2 dataset).

5.7.2 AABP-SUB

The AABP-SUB model was then applied to merge the tm-erdas2 and geovantage_ratio2 datasets. The same window size of 9x9 and thresholds θ=-1, θ=0.2,

www.manaraa.com

θ=0.4, and θ=0.8 were used. The corresponding pan sharpened images are sharpbilin234-1rdwts-db2-sbm-aabp-9--1-whl-nohm.tif, sharpbilin234-1rdwts-db2-sbm-aabp-9-0.4-whl-nohm.tif, and sharpbilin234-1rdwts-db2-sbm-aabp-9-0.8-whl-nohm.tif in the tm-erdas2 folder. The pan sharpened images for the geovantage_ratio2 dataset are - sharp-1rdwts-db2-sbm-aabp-9--1-whl-nohm.tif, sharp-1rdwts-db2-sbm-aabp-9-0.4-whl-nohm.tif, and sharp-1rdwts-db2-sbm-aabp-9-0.8-whl-nohm.tif. Though these images retain the spectral information of the original MS image, they suffer from poor spatial quality. The images appear blurry compared to the PAN image or the pan sharpened image created using two decomposition levels and the substitutive method for this dataset.

5.7.3 Summary of the AABP Model

Both the additive and the substitutive AABP models resulted in poor spatial quality, the additive model was the worse of the two. This is highly contradictory to the published results of the authors and requires further investigation. Since this model did not give good results there is no point in comparing its spectral quality with that provided by the additive or substitutive methods studied earlier.

5.8 Comparisons between the RWT, AWT and the LP

In this section the spectral quality of the RWT, AWT, and LP pan sharpened images are compared in order to decide which one gives the highest spectral fidelity. The comparison of metrics is restricted to RMSE(B^* | F) for simplicity as the use of Wald's property 1 metrics also lead to similar conclusions. The 'db2' wavelet is used for the

www.manaraa.com

RWT, the Gaussian filter corresponding to a=0.5 for the AWT, and the 13-tap QMF for the LP.

Table 5-17 Some properties of the RWT, AWT, and the LP

	RWT	AWT	LP
Directional selectivity	3	none	none
Decimation	none	none	The approximation images are subsampled by a factor of 2 at each scale
Filter	Daubechies, Biorthogonal wavelets	The AWT and LP use the same filters but these are different from those used in the RWT or DWT for e.g. Gaussian, binomial, QMF.	

5.8.1 Comparisons under the NULL Rule

Table 5-18 gives the spectral fidelity metrics of the three transforms under the NULL rule. The three values in each cell denote the metrics for the three bands; this convention is used throughout the remainder of this chapter. It is seen that the spectral fidelity of the RWT is the highest, the LP is the second best and its spectral fidelity is quite close to the RWT. The AWT provides the least spectral fidelity compared to the LP and the RWT. Table 5-19 gives the relative improvement in the spectral fidelity for the RWT and LP with the metrics of the AWT as the reference. The improvement for the tm-vorarlberg2 dataset is the highest, around 20% while for the other three datasets it is somewhat less around 10-15%.

www.manaraa.com

Table 5-18 Spectral quality metrics of RWT, AWT, and LP under NULL rule

Dataset	LP			AWT			RWT		
tm-erdas2	2.507	4.623	4.963	2.818	5.253	5.363	2.438	4.465	4.834
tm-vorarlberg2	7.449	7.081	8.515	9.539	8.749	10.775	7.439	7.034	8.305
qb1	3709	3740	3814	4206	4145	4192	3599	3633	3711
qb-pyramids2	3572	3503	3698	3994	3833	4031	3487	3417	3608

Table 5-19 Relative improvement of LP and RWT over the AWT under NULL rule

Dataset	LP (%)			RWT (%)		
tm-erdas2	11.04	11.99	7.46	13.48	15	9.86
tm-vorarlberg2	21.91	19.07	20.97	22.01	19.6	22.92
qb1	11.82	9.77	9.02	14.43	12.35	11.47
qb-pyramids2	10.57	8.61	8.26	12.69	10.85	10.49

5.8.2 Comparisons under the MAS Rule

Next the spectral fidelity provided by these transforms under the MAS rule is compared. The relative improvements due to the LP and RWT over the AWT are given in Table 5-20 for the tm-erdas2, tm-vorarlberg2, and the qb1 dataset. It is observed that the improvement in spectral fidelity provided by the RWT over the AWT under the MAS rule is significant. Although the improvement of the LP over the AWT is good for the tm-vorarlberg2 dataset, it is very small for the qb1 dataset and even negative for the tm-

www.manaraa.com

erdas2 dataset. The main point to be noted from the comparisons made in Table 5-19 and Table 5-20 is that the relative improvements of LP and RWT with respect to the AWT have reduced by applying the MAS rule. It is seen that the AWT based pan sharpening benefits more from the use of the MAS rule compared to the NULL rule.

Table 5-20 Relative improvement of LP and RWT over the AWT under MAS rule

Dataset	LP (%)			RWT (%)		
tm-erdas2	-0.62	1.57	-3.51	7.4	9.03	2.56
tm-vorarlberg2	13.76	13.15	15.12	16.43	17.06	21.18
qb1	3.17	2.49	2.43	8.59	7.35	6.5

Thus the main conclusion to be drawn from this analysis is that the RWT provides the highest spectral fidelity compared to the AWT or the LP.

5.9 Effect of Directional Selectivity

The Steerable Pyramid Transform (SPT) was used to pan sharpen the datasets. In this experiment the number of orientation bands in the transform is varied between one and six and the MAS rule is applied to choose the high frequency coefficients in order to understand the effect of the directional selectivity. For each dataset, the other parameters of the pan sharpening process like the application of the histogram match and the number of decomposition levels used are kept the same as when the RWT or the LP was used to process it.

www.manaraa.com

Table 5-21 gives the spectral quality metrics of the tm-erdas2 dataset. The first column shows the metrics for the case when the NULL selection rule is applied. In this case varying the number of oriented bands does not affect the output results, since all the high frequency coefficients come from the PAN image.

Table 5-21 Spectral quality metrics for the SPT using K orientation bands (tm-erdas2)

	NULL	K=1	K=2	K=3	K=4	K=6
$Corr(F_1, F_2)$	0.972	0.971	0.971	0.971	0.971	0.971
$Corr(F_2, F_3)$	0.648	0.643	0.643	0.643	0.643	0.643
$Corr(F_3, F_1)$	0.654	0.649	0.649	0.649	0.649	0.649
$Corr(B_1, B_2)$	0.967					
$Corr(B_2, B_3)$	0.601					
$Corr(B_3, B_1)$	0.607					
$Corr(B_1^*, F_1)$	0.965	0.972	0.972	0.972	0.972	0.972
$Corr(B_2^*, F_2)$	0.968	0.972	0.973	0.973	0.973	0.973
$Corr(B_3^*, F_3)$	0.968	0.972	0.972	0.972	0.972	0.972
$RMSE(B_1^*, F_1)$	2.811	2.562	2.557	2.551	2.551	2.550
$RMSE(B_2^*, F_2)$	5.219	4.885	4.876	4.868	4.868	4.865
$RMSE(B_3^*, F_3)$	5.43	5.235	5.231	5.226	5.226	5.225

From the above table it is seen that as the number of oriented bands is increased from one to three there is improvement in the spectral fidelity. However, when K is increased beyond three there is very little improvement. Surprisingly, the inter band correlation does not change much as the number of orientation bands are changed. The relative improvement in the spectral fidelity as the oriented bands are increased is given in Table 5-22, which shows the improvement in percentage as K is increased successively. These metrics show that there is practically no point in using more than three oriented bands.

www.manaraa.com

Table 5-22 Relative improvement in spectral fidelity as a function of K (tm-erdas2)

Percentage Improvement in	NULL→K=1	K=1 → K=2	K=2 → K=3	K=3 → K=4	K=4 → K=6
$RMSE(B_1^*, F_1)$	8.858	0.195	0.234	0	0.039
$RMSE(B_2^*, F_2)$	6.40	0.184	0.164	0	0.062
$RMSE(B_3^*, F_3)$	3.59	0.076	0.096	0	0.019

The same experiment is repeated on the qb1 dataset. The other parameters of the pan sharpening process are as given previously – three decomposition levels, single band processing, and performing the histogram match. Table 5-23 gives the spectral quality metrics for this dataset and Table 5-24 gives the relative improvement in percentage as the number of oriented bands of the SPT is increased. Similar to the previous dataset the improvement obtained by using just K=1 and MAS rule over the NULL rule is significant. There is hardly any improvement when K is increased beyond three. The inter band correlation slightly improves as K is increased to 2 but remains constant thereafter as the number of orientation bands are increased further.

www.manaraa.com

Table 5-23 Spectral quality metrics for the SPT using K orientation bands (qb1)

	NULL	K=1	K=2	K=3	K=4	K=6
$Corr(F_1, F_2)$	0.989	0.987	0.988	0.988	0.988	0.988
$Corr(F_2, F_3)$	0.991	0.989	0.990	0.990	0.990	0.990
$Corr(F_3, F_1)$	0.964	0.960	0.961	0.961	0.961	0.961
$Corr(B_1, B_2)$	0.983					
$Corr(B_2, B_3)$	0.986					
$Corr(B_3, B_1)$	0.950					
$Corr(B_1^*, F_1)$	0.895	0.924	0.925	0.925	0.925	0.925
$Corr(B_2^*, F_2)$	0.902	0.928	0.929	0.929	0.929	0.929
$Corr(B_3^*, F_3)$	0.908	0.929	0.930	0.930	0.930	0.930
$RMSE(B_1^*, F_1)$	3787	3295	3274	3270	3261	3256
$RMSE(B_2^*, F_2)$	3755	3282	3269	3262	3257	3255
$RMSE(B_3^*, F_3)$	3859	3436	3420	3413	3415	3410

Table 5-24 Relative improvement in spectral fidelity as a function of K (qb1)

Percentage Improvement in	NULL→K=1	K=1 → K=2	K=2 → K=3	K=3 → K=4	K=4 → K=6
$RMSE(B_1^*, F_1)$	12.99	0.64	0.12	0.28	0.15
$RMSE(B_2^*, F_2)$	12.6	0.40	0.21	0.15	0.06
$RMSE(B_3^*, F_3)$	10.97	0.47	0.20	-0.06	0.15

If practical considerations are taken into account, an improvement of the order less than 1% would be considered ineffective. Then K=1 seems to be sufficient for pan sharpening when the SPT is used. However, there is definitely a small improvement in the spectral fidelity as the number of orientation bands is increased. This experiment

www.manaraa.com

proves that higher directional selectivity results in little improvement in the spectral fidelity. The findings from the tm-vorarlberg2 dataset are also similar, the improvement in the spectra fidelity when K is increased beyond two are less than 1%.

The main purpose of these experiments was to understand the role that directional selectivity played in improving the performance of RWT over the LP or the AWT. The RWT has three detail images in three directions – H, V, D and the LP has one detail image. Thus the effect of incrementing the number of orientation bands from K=1 to K=3 must be understood in order to make a fair evaluation. Table 5-25 shows the spectral quality improvement obtained by incrementing the number of orientation bands from K=1 to K=3. The improvement is quite small, on the order of 1%. The spectral fidelity of the RWT was better than the LP by about 2-3%, thus it can be concluded that directional selectivity is only partly responsible for the superior performance of the RWT over the LP or the AWT. The choice of the mother wavelet is a more critical factor in determining the spectral fidelity of the pan sharpened image. For example, if the RWT and LP were compared by using the 'db20' wavelet and the 13-tap QMF respectively, the percentage improvement due to the RWT would have been much less. Similarly if a Gaussian filter were used for the LP and the 'db2' wavelet for RWT in order to make comparisons the relative improvement due to the RWT over LP would be much higher.

www.manaraa.com

Table 5-25 Spectral quality improvement in the SPT method (K=1 to K=3)

Percentage Improvement in	tm-erdas2	qb1
$RMSE(B_1^*, F_1)$	0.43	1.06
$RMSE(B_2^*, F_2)$	0.35	0.76
$RMSE(B_3^*, F_3)$	0.17	0.61

5.10 Estimation of the Minimum Number of Decomposition Levels

In order to estimate the minimum number of decomposition levels in MRA based pan sharpening for a particular resolution ratio of the images, experiments were done on datasets created from the 'geovantage' imagery. Four different resolution ratios n=2, 3, 4, and 5 were used. Since the higher resolution MS image is available for comparison for these datasets (the original aerial photograph) the task of analyzing the pan sharpened images is simplified somewhat.

In order to see if the findings from the above datasets can be extended to MS-PAN image pairs where the higher resolution MS image is not available for comparison, the tm-erdas2 and qb1 datasets are also studied. The following sections will describe the results obtained for each of the datasets.

5.10.1 Dataset – geovantage_ratio2

This dataset has a resolution ratio of two. The images were pan sharpened with the RWT based substitutive method. Nine decomposition levels – 1, 2, 3 … 9 were used,

www.manaraa.com

thus nine pan sharpened images are created. Hypothetically an image of size NxN can be decomposed up to at most $\log_2 N$ levels, 10 decomposition levels are the maximum for a 600x600 image (for image sizes not a power of two the next higher power of two integer is considered as N). Since the images are 600x600, they can be decomposed ten times. However, after 7 or 8 levels the computation becomes excessive and thus, decomposition levels up to nine are only considered. Also, the difference is negligible between 9 and 10 decomposition levels results. The pan sharpened images are compared to the higher resolution original MS image by using Wald's property 2. The results for this are given in Table 5-26. The three values in each cell are for the three bands.

Table 5-26 Wald's property 2 and TE metrics (geovantage_ratio2)

L	RMSE	Correlation	True Edges (%)
1	15, 14, 16	0.96 0.95 0.95	35 43 34
2	12 11 13	0.98 0.97 0.97	68 68 64
3	8.8 8.5 10	0.99 0.98 0.98	80 75 71
4	7.7 7.3 9.1	0.99 0.99 0.99	83 76 72
5	7.9 7.3 9.0	0.99 0.99 0.99	84 76 72
6	8.6 7.7 9.3	0.99 0.99 0.98	84 76 72
7	9.4 8.3 9.8	0.99 0.98 0.98	85 76 73
8	10 8.7 10	0.98 0.98 0.98	85 76 73
9	10 8.8 10	0.98 0.98 0.98	85 76 73

www.manaraa.com

It is seen that pan sharpened images with L=1 have very high RMSE and also a poor spatial quality. As L (decomposition levels) is increased up to 4 the RMSE decreases and the correlation also improves. Thus, by looking at Wald's property 2 metrics L=4 seems to be the optimal choice and if L is increased beyond 4 the spectral quality starts degrading. Thus L=4 can be considered the optimal decomposition level for images with n=2.

The pan sharpened images were also visually compared with the original image. The spatial quality of the pan sharpened image created by L=1 was not very good, it appeared blurry. Some of the small structures on the rooftops of a building that were separated in the original image appeared joined in this image. These are shown in Figure 5-11.

Pan sharpened images created with L=2 or higher were much better, and the structures appear disjoint as in the original image. This is also seen from the significant increase in the number of edges extracted from the image. No changes can be noted visually in the pan sharpened images created with L=2 and higher. A comparison of the edge metrics of the pan sharpened images created by using L=2 and 3 respectively shows that significantly more edges have been retained in the pan sharpened image created by using L=3. Thus, if the user is satisfied with modest spatial quality and if computational power and achieving the highest spectral fidelity is an important criterion, L=2 would be a better choice otherwise according to the edge metrics L=3 will give sharper images with better spatial quality. The increase in the edge metrics beyond L=3 is little.

www.manaraa.com

(a) The original MS image

(b) Pan sharpened image, L=1

(c) Pan sharpened image, L=2

(d) Pan sharpened image, L=3

Figure 5-11 The original and various pan sharpened images for dataset I

5.10.2 Dataset – tm-erdas2

The second dataset is the actual TM MS-PAN image pari. Images were pan sharpened using nine different decomposition levels: L=1, 2, 3… 9. The results for the pan sharpened images are presented in Table 5-27. In order to conserve space the metrics for L=6, 7, 8 are not presented here onwards. These values lie between those of L=5 and L=9.

www.manaraa.com

Table 5-27 Spectral quality and TE metrics (tm-erdas2)

L	1	2	3	4	5	9
$Corr(B_1(l) \mid F_1'(l))$	1.0	0.99	0.98	0.97	0.96	0.96
$Corr(B_2(l) \mid F_2'(l))$	1.0	0.99	0.99	0.99	0.98	0.98
$Corr(B_3(l) \mid F_3'(l))$	1.0	0.99	0.99	0.98	0.97	0.95
$Corr(F_1, F_2)$	0.97	0.96	0.96	0.97	0.97	0.97
$Corr(F_2, F_3)$	0.61	0.62	0.64	0.66	0.69	0.72
$Corr(F_3, F_1)$	0.62	0.63	0.65	0.66	0.67	0.68
$Corr(B_1, B_2)$	**0.97**					
$Corr(B_2, B_3)$	**0.60**					
$Corr(B_3, B_1)$	**0.61**					
$Corr(B_1^*, F_1)$	0.98	0.96	0.95	0.94	0.94	0.93
$Corr(B_2^*, F_2)$	0.99	0.98	0.98	0.97	0.97	0.97
$Corr(B_3^*, F_3)$	0.99	0.98	0.97	0.97	0.96	0.94
True Edges (%)	57	63	63	62	62	61
	58	66	66	65	65	65
	58	71	73	74	74	75

Visual analysis showed that the pan sharpened image produced by keeping L=1 is quite blurry in certain regions, which is also reflected in the edge metric. The edge metric is more or less constant after L is increased beyond 2. Thus the optimal choice would be L=2 since the spectral quality of the image does not degrade much compared to L=1. In this case a higher resolution truth image is not available, and thus the choice of the optimal decomposition level is difficult. The choice must be made based on both the spectral and spatial quality metrics. It is seen that as the decomposition levels increase, the spatial quality improves and saturates after a certain number of decompositions. While at the same time it is known that the spectral similarity with respect to the low resolution MS images decreases. According to Wald's property 2 the pan sharpened

www.manaraa.com

images are approaching the ideal high resolution MS image as L is increased to 3 or 4, which is good. However, the spectral similarity with respect to the original MS images decreases.

Thus, it is seen that a different choice of metrics gives a different optimal decomposition level. It is recommended that if the pan sharpened images are to be used for automated classification where the spectral signatures from the original low resolution MS images will be used, then the lower decomposition levels are recommended (L=2). This is because it preserves the radiometry of the original image well and delivers sufficient spatial quality at the same time.

5.10.3 Dataset – geovantage_ratio3

This dataset has a resolution ratio of three. The spectral and spatial quality metrics of the various pan sharpened images are given in Table 5-28.

Table 5-28 Wald's property 2 and TE metrics (geovantage_ratio3)

L	RMSE	Correlation	True Edges (%)
1	12 12 13	0.97 0.97 0.97	56 53 61
2	8.9 9 11	0.99 0.98 0.98	81 76 73
3	7.3 7.2 8.9	0.99 0.99 0.99	88 79 77
4	7 6.9 8.5	0.99 0.99 0.99	90 79 77
5	7.5 7.2 8.7	0.99 0.99 0.99	90 79 77
9	9.8 8.8 10	0.99 0.98 0.98	91 79 77

www.manaraa.com

A choice of L=4 gives the minimum RMSE and the highest possible spatial quality. The correlation and RMSE start deteriorating for L=5 or higher, thus the metrics for L=6, 7, 8 are not given. However, the RMSE values for L=3 are not significantly different from those of L=4. Thus, L=3 or 4 would be an optimal choice for pan sharpening images with a resolution ratio of 3.

Pan sharpened images produced from L=1 were blurry and the shape of many structures could not be reconstructed correctly. Many features in the pan sharpened image appeared smeared and two nearby objects appeared joined when in fact they are separated by a small distance. Pan sharpened images produced from L=2 appeared to be satisfactorily merged. They were sharp and all the details present in the original color photograph were evident. The edge metrics shows a higher amount of improvement in first band when L is increased from 2 to 3 while the improvement for the other two bands is lesser and quite insignificant.

A literature search revealed that Lemeshewsky [18] also used a RWT with three decomposition levels to merge images with n=3. The images were a TM MS (30 m) and a SPOT PAN (10 m).

5.10.4 Dataset – geovantage_ratio4

This dataset is a MS-PAN image pair with ratio four. The spectral and spatial quality metrics of the various pan sharpened images are given in Table 5-29.

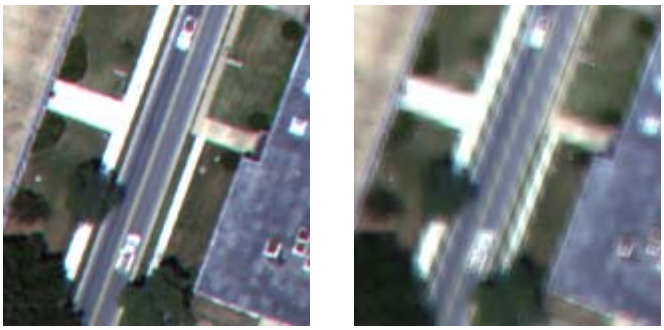
www.manaraa.com

Table 5-29 Wald's property 2 and TE metrics (geovantage_ratio4)

L	RMSE	Correlation	True Edges (%)
1	18 17 19	0.94 0.93 0.94	44 44 43
2	14 14 16	0.96 0.96 0.96	75 75 74
3	10 9.9 12	0.98 0.98 0.97	88 85 83
4	8.1 8 10	0.99 0.99 0.98	91 85 84
5	8 7.7 9.7	0.99 0.99 0.98	92 85 84
9	10 9.1 11	0.98 0.98 0.98	93 85 83

In this case a choice of L=4 or 5 seems the optimal choice. The difference in terms of RMSE between L=4 or 5 is insignificant, thus practically L=4 is a better choice. The correlation and RMSE start deteriorating for L=6 or higher. As noted previously, pan sharpened images from L=1 were blurry. Though the images from L=2 are not blurry, they too have a smearing effect in some parts of the image. This is shown in Figure 5-12 where it is seen that the road appears slightly unmerged. Thus, visually a value of L=3 seems sufficient to pan sharpen images with a resolution ratio of 4. If the edge metrics are used to determine the minimum 'L' they also predict that a value of L=3 is sufficient as the improvement in edge metrics beyond L=3 is insignificant.

www.manaraa.com



(a) The original MS image (b) Pan sharpened image, L=1

(c) Pan sharpened image, L=2 (d) Pan sharpened image, L=3

Figure 5-12 The original and pan sharpened images for dataset IV

5.10.5 Dataset – qb1

The fifth dataset is a Quickbird MS-PAN pair, which also has a resolution ratio of 4. The metrics for the pan sharpened images of this dataset are given in Table 5-30.

www.manaraa.com

Table 5-30 Spectral quality and TE metrics (qb1)

L	1	2	3	4	5	9
$Corr(B_1(l) \mid F_1'(l))$ $Corr(B_2(l) \mid F_2'(l))$ $Corr(B_3(l) \mid F_3'(l))$	0.96 0.96 0.97	0.96 0.96 0.97	0.94 0.94 0.94	0.91 0.9 0.91	0.90 0.89 0.90	0.88 0.85 0.87
$Corr(F_1, F_2)$ $Corr(F_2, F_3)$ $Corr(F_3, F_1)$	0.98 0.99 0.95	0.98 0.99 0.95	0.98 0.99 0.95	0.98 0.99 0.96	0.98 0.99 0.95	0.98 0.99 0.95
$Corr(B_1, B_2)$ $Corr(B_2, B_3)$ $Corr(B_3, B_1)$	0.98 0.99 0.95					
$Corr(B_1^*, F_1)$ $Corr(B_2^*, F_2)$ $Corr(B_3^*, F_3)$	0.98 0.08 0.98	0.94 0.93 0.94	0.91 0.90 0.91	0.9 0.88 0.89	0.89 0.87 0.89	0.87 0.84 0.86
True Edges (%)	34 38 40	69 78 81	78 89 91	80 93 93	76 85 85	76 85 85

It was clear from visual inspection of the fused images that one or two levels of decomposition were insufficient. Visual analysis of the fused images from three decomposition levels onwards confirms that the spatial information is fused satisfactorily. They appear identical. The increase in the edge metrics when L is increased from 3 to 4 is

www.manaraa.com

not much. Strangely, when L is increased beyond four, the number of edges extracted also decrease. Thus, there is no advantage in going beyond a certain number of decomposition levels. In fact, as the spectral quality metrics show it is a disadvantage.

Similar to the observations made for dataset II, it is seen that as the decomposition levels increase the spectral similarity between the original MS and the sharpened images decreases. At the same time the amount of correlation between the sharpened and PAN image increases. The correlation between the original MS and PAN bands was 0.79, 0.84 and 0.85 and that between the PAN and the sharpened bands from 3 decomposition levels was 0.9, 0.94 and 0.95. While for 9 decomposition levels it is 0.95, 0.99 and 0.99. Thus, it can be concluded that as the decomposition levels increase, a higher percentage of information from the PAN image is retained in the fused image and less from the MS image.

Therefore, it is shown that more than four decomposition levels are redundant and the choice must be made between 3 and 4 levels. The lower bound preserves the radiometry of the MS images better while 4 decomposition levels leads to slightly better spatial quality (i.e., the images look sharp and visually pleasing). Preserving the original radiometry of the images is important from the point of view of automated classification while extracting the maximum edge information from the PAN image is important for tasks like automated object extraction (e.g. buildings, roads). Thus, the number of decomposition levels must be chosen depending upon the application.

www.manaraa.com

5.10.6 Dataset – geovantage_ratio5

The final dataset is a MS-PAN pair with ratio five. The metrics for the pan sharpened images are given in Table 5-31.

Table 5-31 Wald's property 2 and TE metrics (geovantage_ratio5)

L	RMSE	Correlation	True Edges (%)
1	18 18 20	0.94 0.92 0.93	46 43 44
2	15 14 16	0.96 0.95 0.95	77 75 75
3	10 10 12	0.98 0.98 0.97	90 87 86
4	8.1 8.1 10	0.99 0.99 0.98	93 88 86
5	8 7.9 10	0.99 0.99 0.98	94 88 86
9	10 9.3 11	0.98 0.98 0.98	95 87 86

It is seen from the metrics that the pan sharpened images created with L=5 is the optimal, i.e. most closely resembles the original high resolution color photograph. The spatial quality is also good. The RMSE and the correlation with the high resolution MS images starts decreasing after L=6. Visually, the images generated by L=3 appeared satisfactorily merged. The number of edges extracted drastically increases as L is changed from 1 to 2 or 2 to 3, but is more or less constant after L is increased beyond that. All the structures observed in the original image could be seen in the pan sharpened image corresponding to L=3.

www.manaraa.com

5.10.7 Summary

The findings from the above experiments are summarized in this subsection. It was seen that there is not a unique answer to the problem of finding the "optimal number of decomposition levels for images with a particular resolution ratio". The decision should be based on the purpose of the pan sharpened image. For example, if the pan sharpened images are to be used for automatic or supervised classification in order to create land use maps the lower number of decomposition levels which better preserve the spectral similarity with the MS images must be applied. For tasks like automated extraction of urban objects (e.g., roads, buildings) which require rich spatial information and preservation of the original spectral information, the higher number of decomposition levels is justified. Since MRA based methods have this characteristic they are preferred compared to computationally simpler statistical methods like IHS or Brovey. Thus, in Table 5-32 we make two recommendations for the optimal decomposition level for each resolution ratio based on the application. In the third column of the table the minimum number of decomposition levels estimated by the TE metric is also given. There are two values in the first row corresponding to n=2 because the estimations for the tm-erdas2 and geovantage_ratio2 datasets are different. The user can choose the number of decomposition levels depending upon the application of the pan sharpened image and the computational power available to them.

www.manaraa.com

Table 5-32 Recommended decomposition levels

Resolution Ratio (n)	Urban feature extraction	Classification for land use mapping	Estimated value by TE
2	4	2	3,2
3	4	2	2
4	5	3	3
5	5	3	3

The TE metric was proposed as an indicator of the minimum number of decomposition levels required to merge the pan sharpened images with sufficient spatial quality as determined by visual analysis. It is seen that except for the first dataset – geovantage_ratio2 the value of 'L' predicted by TE was in agreement with the number deemed sufficient by visual analysis. In section 6.2.1 it is suggested how this metric could be used to predict the minimum number of decomposition levels required to merge images whose resolution ratio may not be known.

www.manaraa.com

CHAPTER VI

CONCLUSIONS AND FUTURE WORK

In this chapter conclusions are drawn based on the experiments performed in this study and the contributions deriving from this work. Also, topics that are worthy of further investigation are suggested for future work.

6.1 Conclusions

6.1.1 Choice of the Wavelet for RWT

Various mother wavelets were evaluated and it was found that the 'db2' or the second order Daubechies wavelet provided the highest spectral fidelity for the RWT based pan sharpening. This wavelet even outperformed the biorthogonal wavelets. The analysis of Daubechies wavelets showed that the shorter wavelets preserved the spectral fidelity to a higher degree compared to the longer wavelets with the exception of the 'db1' (Haar) wavelet. The spectral fidelity of this wavelet was poorer than the longest wavelets that were used.

6.1.2 Choice of the Filter for LP and AWT

It was discovered that for the LP based pan sharpening the 13-tap QMF gave the highest spectral fidelity. Three different families of filters were used – the Gaussian

www.manaraa.com

filters, the binomial filters and QMFs. A general relationship between the characteristics of the filters with the spectral fidelity was determined. It was discovered that as the length of the QMFs increased the spectral fidelity improved while for the binomial filters as the filter length increases the spectral fidelity degrades. For the Gaussian filter as the weight of the central coefficient of the filter is increased the spectral fidelity improves. However, when the filter becomes trimodal (i.e. the weight of the central coefficient is 0.6 or more) the pan sharpened image contains artifacts and the spatial quality is not acceptable.

The same filters were used to perform AWT based pan sharpening, and it was determined that Gaussian filters gave higher spectral fidelity compared to the binomial filters or QMFs. The literature survey showed that for AWT based pan sharpening almost all researchers used the 'cubic spline' filter or the length 5 Gaussian filter with the central coefficient weight equal to 0.375. I have shown here that better results can be obtained with the AWT using a different filter. It was seen that as the weight of the central coefficient of the Gaussian filter was increased up to 0,5 the spectral fidelity improved but degraded drastically for higher values.

6.1.3 Choice of the MRA Transform

Three MRA transforms – RWT, LP, and AWT were compared and it was found that the RWT gave the highest spectral fidelity. The spectral fidelity of the LP is only slightly less than that of the RWT while that of the AWT is much poorer.

To the best of my knowledge a detailed comparison of the three transforms has not been made. The novelty of this approach is that before arriving at the final conclusions the best wavelet (or filter) is used for each transform.

www.manaraa.com

6.1.4 Comparison of the Additive and Substitutive Methods

To the best of my knowledge a detailed comparison between the additive and substitutive methods was not done previously. It was found that the additive method based on the RWT is not suitable, the images pan sharpened with this method have very poor spectral fidelity. The additive method based on the AWT delivers good results when pan sharpening images with resolution ratio of two, the spectral quality is higher than that of the substitutive RWT method, although for higher ratios (e.g., four) it results in poor spatial quality. This discovery that the additive method based on the AWT results in poor spatial quality is in contradiction with the findings of Nunez et al. [26]. They found this method to be better than the substitutive method based on the AWT and did not find any problems with the spatial quality of the pan sharpened images.

The LP based additive method provides high spectral quality comparable to that of the substitutive methods. One of the main advantages of the LP based additive method or the additive methods in general is that only the PAN image has to be decomposed. Thus, it will be computationally cheaper compared to the substitutive method in which both the PAN and MS images must be decomposed.

However, the drawback of the additive method is that sometimes it provides slightly less spatial quality compared to the substitutive methods, and it can only be applied to pan sharpen MS-PAN images whose resolution ratio is a power of two.

6.1.5 Choice of the Selection Rule

The various selection rules available to combine the detail coefficients of the PAN and the MS image were studied and it was determined that the point wise maximum

www.manaraa.com

amplitude selection (MAS) rule produces images with the highest spectral fidelity compared to the window based methods. In summary the three main selection rules can be ranked in the order 1) MAS, 2) WBS, and 3) WBV based on the spectral fidelity criterion. Thus the MAS rule must be preferred because not only does it result in the highest spectral fidelity but it is also computationally much cheaper compared to the WBS and WBV rules. However, it must be understood that sometimes these rules can cause slightly pixilated edges in some regions of the image and slight smearing around sharp objects. These artifacts are relatively few and do not always occur, moreover they are only noticeable to a keen observer.

To the best of my knowledge a comparison of these selection rules in the context of pan sharpening was not done previously.

Thus, if the images with the highest spectral fidelity possible are desired, the RWT based substitutive method is recommended, and a Daubechies 2nd order wavelet must be used. The maximum amplitude selection (MAS) rule must be applied to choose between the high frequency coefficients of both images.

6.1.6 The Effect of the Directional Selectivity

There is a misconception in the image fusion community that the reason for the superior performance of the RWT with respect to the LP is its higher directional selectivity. Through experiments on the SPT, I have demonstrated that the higher directional selectivity plays only a minor role in improving the spectral fidelity of the pan sharpened image. The major reason for the higher spectral fidelity provided by the RWT with respect to the LP or the AWT is the fact that a better wavelet is used compared to

www.manaraa.com

the filter used for the LP or the AWT. If an inferior wavelet were used for the RWT like the 'db20' the relative spectral fidelity improvement from the use of the RWT over the LP could be lower.

6.1.7 Estimation of the Minimum Number of Decomposition Levels

The quality of the pan sharpened images produced by the wavelet based fusion algorithm depends upon the number of decomposition levels in the MRA transform. Too few decomposition levels result in pan sharpened images with poor spatial quality, while excessive levels severely distort the radiometry of the pan sharpened images compared to the original MS images. The wavelet transform has a high computational complexity that increases with the number of decomposition levels. Thus determining the minimum number of decomposition levels is important. The minimum number of decomposition levels was estimated by conducting a systematic study using both objective and subjective metrics. This study determined the minimum number of decomposition levels necessary in the wavelet based pan sharpening for PAN-MS image pairs with different resolution ratios. The larger the resolution ratio is, the more decomposition levels are required to achieve a satisfactory fusion result.

One interesting fact derived from the study was that a different choice of metrics leads to a different recommendation on the minimum number of decomposition levels. When a truth or reference image is available for comparison, the minimum decomposition level is chosen based on the pan sharpened image which most closely resembles the reference image in terms of pixel values. In other words this image is the closest that a higher resolution MS sensor would have observed if it existed. While if a

www.manaraa.com

reference image is not available the minimum number is more difficult to determine or interpret. It is chosen as the image that achieves sufficient spatial quality compared to the pan sharpened images and also preserves the radiometry of the unmerged MS images as much as possible. However, these are conflicting goals as seen from the results. The spectral similarity between the original MS and the pan sharpened images decreases as the decomposition levels increase. While the spatial quality of the pan sharpened images increases up to a certain number of decompositions but remains constant after that.

For example, when working on image pairs with a resolution ratio of two, comparison with the reference image indicated that L=4 was the best choice while in its absence L=2 seemed a better choice. Similarly for image pairs with resolution ratio four, the minimum decomposition levels as determined by the two different methods were 5 and 3. The difference in the two decisions is quite considerable. Four decomposition levels would take up a lot more computation than two since remote sensing datasets are generally huge. Thus it is advocated that when it is necessary to preserve the radiometry of the original MS images in the pan sharpened images, the minimum required decomposition levels be applied. This is useful for tasks like automated classification, when sometimes the sharpened imagery is classified using spectral signatures from the original MS image. However, if the application is automated object extraction (e.g. buildings, roads) where images with high spatial and high spectral information are desired the higher decomposition levels are recommended.

www.manaraa.com

Recommendations for the minimum number of decomposition levels required to merge images with resolution ratios of 2, 3, 4 and 5 are given in section 5.10.7 based on the application of the pan sharpened image.

6.2 Potential Topics for Future Work

Finally, this section describes an interesting topic for future work that could not be carried out due to time limitation.

6.2.1 Automatic Estimation of the Number of Decomposition Levels

Suppose it is decided to make available the MRA based pan sharpening as part of a remote sensing software so that users can pan sharpen their images without any effort by simply specifying the MS and PAN images to the software. The parameters of the pan sharpening like the wavelet or the filter used and the MRA transform can be kept fixed in the software since the best one has been determined through empirical evaluations. For example if the substitutive method is used the RWT and the 'db2' wavelet are recommended. The histogram match option can also be automatically decided by computing the correlation between the MS and the PAN images.

However, the number of decomposition levels in the MRA based pan sharpening process cannot be kept fixed since it depends on the resolution ratio of the images. The simplest thing to do would be to let the user supply this information. Many times the organization selling the imagery supplies the MS imagery resampled to the pixel size of the PAN image. In such a scenario there is no way to tell the resolution ratio unless this is specified in the metadata. Even if the resolution ratio is known it is best to make the

www.manaraa.com

software user supply as little information to the algorithm as possible, i.e. the software should have intelligent defaults. Thus, it would be sophisticated to not have the user specify the decomposition levels and the software estimate it based on calculations of the image characteristics.

Any commercial or real software for wavelet based pan sharpening will do block processing since the images are generally huge. The idea is to block process the first few blocks with different decomposition levels (1, 2 … 5) and calculate the True Edge (TE) metric for each pan sharpened block corresponding to the different number of decomposition levels used. It was seen from our experiments that when the spatial quality of the pan sharpened image is satisfactory the TE metric does not improve significantly.

Thus, among the various decomposition levels applied for each block the decomposition level 'L' after which the TE metric changes relatively little (less than 5-10 %) should be chosen as the minimum decomposition level. The prediction for each block may differ by one or two. In such situations the maximum value of 'L' obtained for the processed blocks can be chosen to compute the final pan sharpened image.

www.manaraa.com

REFERENCES

1. Aiazzi B., Alparone L., Barducci A., Baronti S., and Pippi I., "Multispectral fusion of multisensor image data by the generalized Laplacian pyramid", Proceedings of the IEEE International Geoscience And Remote Sensing Symposium, pp. 1183-1185, 1999.

2. Aiazzi B., Alparone L., Baronti S., and Garzelli A., "Context-Driven Fusion of High Spatial and Spectral Resolution Images Based on Oversampled Multiresolution Analysis", vol. 40, no. 10, pp. 2300-2312, 2002.

3. Blanc P., Blu T., Ranchin T., Wald T., and Aloisi T., "Using iterated rational filter banks within the ARSIS concept for producing 10 m LANDSAT multispectral images", International Journal of Remote Sensing, vol. 19, no. 12, pp. 2331-2343, 1998.

4. Burrus C.S., Gopinath R.A., and Guo H., Introduction to Wavelets and Wavelet Transforms – A Primer, Prentice Hall, NJ, 1998.

5. Burt P.J., and Adelson E.H., "The Laplacian Pyramid as a Compact Image Code", IEEE Transactions on Communications, vol. 31, no. 4, pp. 532-540, 1993.

6. Burt P.J., and Lolczynski R.J., "Enhanced image capture through fusion", Proceedings of the Fourth International Conference on Computer Vision, Berlin, Germany, pp. 173-182, May, 1993.

7. Carper W.J., Lillesand T.M., and Kiefer R.W., "The Use of Intensity-Hue-Saturation Transformations for Merging SPOT Panchromatic and Multispectral Image Data", Photogrammetric Engineering and Remote Sensing, vol. 56, no. 4, pp. 459-467, 1990.

8. Castleman K.R., "Digital Image Processing", Prentice Hall Inc., 1996.

9. Chavez P.S. Jr., Sides S.C., Anderson J.A., "Comparison of three different methods to merge multiresolution and Multispectral data: LANDSAT TM and SPOT Panchromatic", Photogrammetric Engineering and Remote Sensing, vol. 57, no. 3, pp. 295-303, 1991.

www.manaraa.com

10. Couloigner L., Ranchin T., Valtonen V.P., and Wald L., "Benefit of the future SPOT 5 and of data fusion to urban mapping", International Journal of Remote Sensing, vol. 19, no. 8, pp. 1519-1532, 1998.

11. Das K., Rao G.K., and Prakash K.A., "Improvement of effective spatial resolution of thermal infrared data for urban landuse classification", Proceedings of the IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Rome, Italy, November 8-9th, pp. 332-336, 2001.

12. ERDAS, "The ERDAS Field Guide" Seventh Version, GIS & Mapping, LLC, Atlanta, Georgia, USA.

13. Fowler J.E., "The Redundant Discrete Wavelet Transform and Additive Noise", IEEE Signal Processing Letters, 2005 (accepted for publication).

14. Gonzalez-Audicana M., Jose L.S., Catalan R.G., and Garcia R., "Fusion of Multispectral and Panchromatic Images Using Improved IHS and PCA Mergers Based on Wavelet Decomposition", IEEE Transactions on Geoscience and Remote Sensing, vol. 42, no. 6, pp. 1291-1299, 2004.

15. Hill P., Canagarajah N., and Bull D., "Image Fusion using Complex Wavelets", Proceedings of the Thirteenth British Machine Vision Conference, 2002.

16. Jain A.K., "Fundamentals of Digital Image Processing", Prentice Hall, 1988, pp. 57-60.

17. King R., and Wang J., "A wavelet based algorithm for pan sharpening Landsat 7 imagery", Proceedings of the International Geoscience and Remote Sensing Symposium, vol. 2, pp. 849-851, 2001.

18. Lemeshewsky G. P., "Multispectral multisensor image fusion using wavelet transforms", Proceedings of the SPIE, vol. 3716, pp. 214-222, 1999.

19. Lemeshewsky G. P., "Multispectral Image Sharpening Using A Shift-Invariant Wavelet Transform and Adaptive Processing of Multiresolution Edges", Proceedings of SPIE, vol. 4736, pp. 189-200, 2002.

20. Li H., Manjunath B.S., and Mitra S.K., "Multisensor Image Fusion Using the Wavelet Transform", Graphical Models and Image Processing, vol. 57, no. 3, pp. 235-245, 1995.

21. Li S., Kwok J.T., and Wang Y., "Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images", Information Fusion 3, pp. 17-23, 2002.

www.manaraa.com

22. Lillesand T., and Kiefer R., Remote Sensing and Image Interpretation, 3rd edition, John Wiley and Sons Inc., 1994.

23. Mallat S.G., "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674-693, 1989.

24. Muhammad S., Wachowicz M., and Carvalho L.M., "Evaluation of Wavelet Transform Algorithms for Multi-Resolution Image Fusion", Proceedings of the Fifth Internation Conference on Information Fusion (International Society for Information Fusion), 2002.

25. Munechika C.K., Warnick J.S., Salvaggio C., and Schott J.R., "Resolution enhancement of multispectral image data to improve classification accuracy", Photogrammetric Engineering and Remote Sensing, vol. 59, pp. 67-72, 1993.

26. Núñez J., Otazu X., Fors O., Prades A., Palà V., and Arbiol R., "Multiresolution-Based Image Fusion with Additive Wavelet Decomposition", IEEE Transactions of Geoscience and Remote Sensing, vol. 37, no. 3, pp. 1204-1211, 1999.

27. Parcharidis L., Kaz-Tani L.M., "Landsat TM and ERS data fusion: a statistical approach evaluation for four different methods", Proceedings of the International Geosciences and Remote Sensing Symposium, vol. 5, pp. 2120-2122, 2000.

28. Pohl C., Genderen J., "Multisensor image fusion in remote sensing: concepts, methods and applications", Internation Journal of Remote Sensing, vol. 19, no. 5, pp. 823-854.

29. Ranchin T., and Wald L., "Fusion of High Spatial and Spectral Resolution Images: The ARSIS Concept and its Implementation", Photogrammetric Engineering and Remote Sensing, vol. 66, no. 1, pp. 49-61, 2000.

30. Ranchin T., Aiazzi B., Alparone L., Baronti S., and Wald L., "Image fusion – the ARSIS concept and some successful implementation schemes", ISPRS Journal of Photogrammetry and Remote Sensing", vol. 58, pp. 4-18, 2003.

31. Remote Sensing Technologies, http://chesapeake.towson.edu/data/tech.asp.

32. Rockinger O., and Fechner T., "Pixel-level Image Fusion: The Case of Image Sequences", Proceedings of the SPIE, vol. 3374, pp. 378-388, 1998.

www.manaraa.com

33. Ryan R., Baldridge B., Schowengerdt R.A., Choi T., Helder D.L., Slawomir B., "IKONOS spatial resolution and image interpretability characteristization", Remote Sensing of Environment, vol. 88, no. 1, pp. 37-52, 2003.

34. Schowengerdt R.A., "Reconstruction of Multispatial, Multispectral Image Data Using Spatial Frequency Content", Photogrammetric Engineering and Remote Sensing, vol. 46, no. 10, pp. 1325-1334.

35. Sharma R., "Probabilistic Model-Based Multisensor Image Fusion", Ph.D. dissertation, Oregon Graduate Institute of Science and Technology, 1999.

36. Shettigara V.K., "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set", Photogrammetric Engineering and Remote Sensing, vol. 58, pp. 561-567, 1992.

37. Shi R., Zhu C., Zhu C., and Yang X., "Multi-Band Wavelet for Fusing SPOT Panchromatic and Multispectral Images", Photogrammetric Engineering and Remote Sensing, vol. 69, no. 5, pp. 513-520, 2003.

38. Simoncelli E.P., Freeman W.T., Adelson E.H., and Heeger D.J., "Shiftable Multiscale Transforms", IEEE Transactions on Information Theory, vol. 38, no. 2, pp. 587-607, 1992.

39. Smith A.R., "Color gamut transform pairs", Proceedings of the 5th annual conference on Computer graphics and interactive techniques, SIGGRAPH, pp. 12-19, 1978.

40. Strang G., and Nguyen T., Wavelets and Filter Banks, 3rd edition, Wellesley-Cambridge Press, Wellesley, MA, 1996.

41. The SPOT website, http://www.spotimage.fr.

42. Vaiopoulos D., Nikolakopoulos K., and Skianis G., "A comparative study of resolution merge techniques and their efficiency in processing image of urban areas", Proceedings of the IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Rome, Italy, November 8-9th, pp. 270-274, 2001.

43. Vijayaraj V., Younan N.H., and O'Hara C.G., "Quality Metrics for Multispectral Image Processing", Proceedings of the Annual ASPRS Conference 2004.

44. Wald L., Ranchin T., and Mangolini M., "Fusion of Satellite Images of Different Spatial Resolutions: Assessing the Quality of Resulting Images", Photogrammetric Engineering & Remote Sensing, vol. 63, no. 6, pp. 691-699, 1997.

www.manaraa.com

45. Yesou H., Besnus Y., Rolet J., and Pion J.C., "Merging Seasat and SPOT imagery for the study of geologic structures in a temperate agricultural region", Remote Sensing of Environment, vol. 43, pp. 265-280, 1993.

46. Yocky D., "Image merging and data fusion by means of the discrete two-dimensional wavelet transform", Journal of Optical Society of America, vol. 12, no. 9, pp. 1834-1841, 1995.

47. Yocky D., "Multiresolution Wavelet Decomposition Image Merger of Landsat Thematic Mapper and SPOT Panchromatic Data", Photogrammetric Engineering & Remote Sensing, vol. 62, no. 9, pp. 1067-1074, 1996.

48. Zhang Y., "A new merging method and its spectral and spatial effects", International Journal of Remote Sensing, vol. 20, no. 10, pp. 2003-2014, 1999.

49. Zhang Y., "Problems in the Fusion of Commercial High-Resolution Satellite as well as LANDSAT 7 Images and Initial Solutions", ISPRS Symposium on Geospatial Theory, Processing, and Applications, vol. 34, no. 4, Ottawa 2002.

50. Zhang Y., and Wang R., "Multi-resolution and multi-spectral image fusion for urban object extraction", Proceedings of the 20th ISPRS Commission 3, pp. 960-966, 2004.

51. Zhang Z., and Blum R.S., "A Categorization of Multiscale-Decomposition-Based Image Fusion Schemes with a Performance Study for a Digital Camera Application", vol. 87, no. 8, pp. 1315-1326, 1999.

52. Zhou J., Civco D.L., and Silander J.A., "A wavelet transform method to merge Landsat TM and SPOT panchromatic data", International Journal of Remote Sensing, vol. 19, no. 4, pp. 743-757, 1998.

www.manaraa.com

APPENDIX A

ACRONYM LIST

www.manaraa.com

AABP	Aiazzi, Alparone, Baronti and Pippi
AABP-ADD	AABP Additive Model
AABP-SUB	AABP Substitutive Model
AMC	Arithmetic Combination
AWT	A Trous Wavelet Transform
COS	Component Based Substitution
DWT	Discrete Wavelet Transform
HM	Histogram Match
HPCC	High Pass Correlation Coefficient
HPF	High Pass Filtering
IHS	Intensity Hue Saturation
IMT	Inverse Multiresolution Transform
LHS	Luminance Hue Saturation
LP	Laplacian Pyramid
MAS	Maximum Amplitude Selection
MRA	Multiresolution Analysis
MT	Forward Multiresolution Transform
PCA	Principal Components Analysis
QMF	Quadrature Mirror Filter
RWT	Redundant Wavelet Transform
SPT	Steerable Pyramid Transform

www.manaraa.com

SVR	Synthetic Variable Ratio
VHS	Value Hue Saturation
WBS	Window Based Salience
WBS	Window Based Verification

www.manaraa.com

APPENDIX B

MRA TRANSFORMS

www.manaraa.com

B.1 The Discrete Wavelet Transform

It is seen in section 2.4.2 that a function can be decomposed into its approximations and details. Thus, a given signal at original resolution J can be decomposed into details at successively lower resolutions J-1, J-2 … J_0 and the approximation at the lowest resolution J_0. This is given by the following equation [4]:

$$f(t) = \sum_k a_{J_0}(k)\phi_{J_0 k}(t) + \sum_k \sum_{j=J_0}^{J} d_j(k)\psi_{j,k}(t), \tag{68}$$

The coefficients $a_{J_0}(k)$ and $d_j(k)$ are the Discrete Wavelet Transform (DWT) coefficients of the function and they are calculated by the inner products of the function with the scaling and wavelet functions respectively. In practice, the DWT coefficients are not calculated in this way. Mallat described a filter bank implementation of the DWT [23] where a set of low pass and high pass filters obtained from the scaling and wavelet functions are convolved with the signal and the output of both the filters are downsampled. The outputs of the low pass and high pass filters correspond to the approximation (low frequency) and detail (high frequency) components of the signal respectively. This procedure can be recursively applied to the output of the low pass filter to obtain successively lower approximations, while the output of the high pass filter at each iteration level is retained. The DWT algorithm can be applied for up to $\log_2 N$ iterations for a signal of length N until we are left with only one sample of the signal.

An image is a two-dimensional signal. Thus, if we want to compute the DWT coefficients of an image, we can apply the DWT algorithm on the rows first and then on the columns. This is possible because the DWT is a separable transform. Figure B-1

www.manaraa.com

illustrates the process of obtaining the DWT of an image. The symbols h_φ and h_ψ are, respectively, the low pass and high pass wavelet filters used for decomposition. In the first step, the rows of the image are convolved with the low pass and high pass filters and the result is downsampled by a factor of two along the columns. The high pass or detailed coefficients characterize the image's high frequency information with vertical orientation while the low pass component contains its low frequency vertical information. Both subimages are again filtered column-wise with the same low pass and high pass filters and downsampled along rows. This DWT produces four subimages at a lower scale – A, H, V, and D. Each subimage is one-fourth the size of the original image. These four components can be interpreted as follows [8]:

1) Approximation coefficients (A) – the intensity or gray-level variations of the image,
2) Horizontal coefficients (H) – variations along the columns,
3) Vertical coefficients (V) – variations along the rows,
4) Diagonal coefficients (D) – variations along the diagonals.

Sequential DWTs can be performed on the approximation image A to yield images at successively lower resolutions.

www.manaraa.com

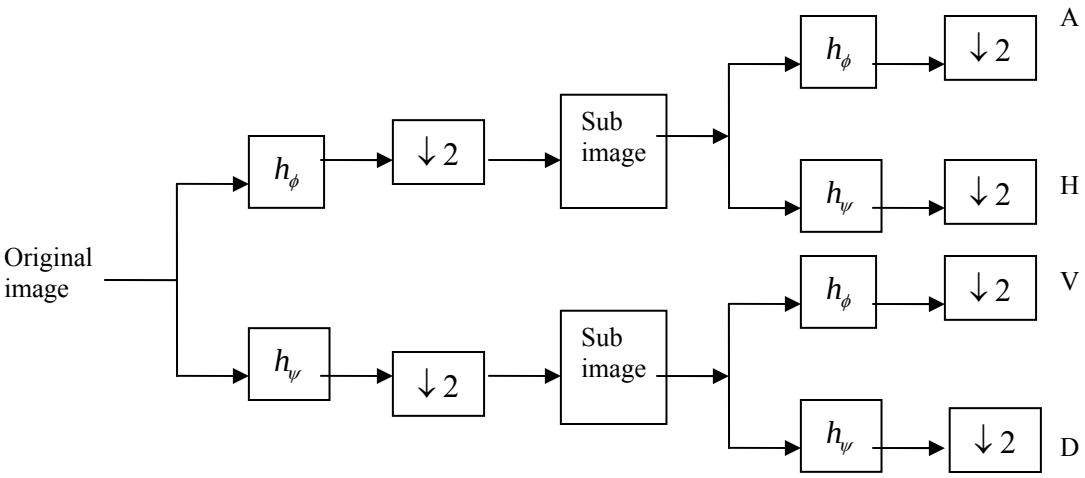


Figure B-1 The 2-D DWT

The Inverse Discrete Wavelet Transform (IDWT) allows us to reconstruct the image at increasingly higher resolutions by using the four subimages, A, H, V, and D obtained from the DWT of the original image. The symbols $\tilde{h}_\varphi$ and $\tilde{h}_\psi$ are the low pass and high pass reconstruction filters respectively. Figure B-2 illustrates the process of reconstructing the original image from its DWT components. Each step is the inverse of a corresponding step in the DWT. The four subimages are upsampled along rows first, since the last step in the DWT was downsampling along rows. The approximation and the vertical subimages are convolved with the low pass filter while the horizontal and detailed subimages are convolved with the high pass filter column-wise (in the DWT we filtered along the columns second). The upsampled and convolved versions of the vertical (V) and diagonal (D) coefficients are summed together, and the upsampled and convolved versions of the approximation (A) and horizontal (H) coefficients are summed

www.manaraa.com

together. Both these intermediate outputs are upsampled along columns and then filtered row wise with the high pass and low pass reconstruction filters, respectively, and summed to yield the original image.

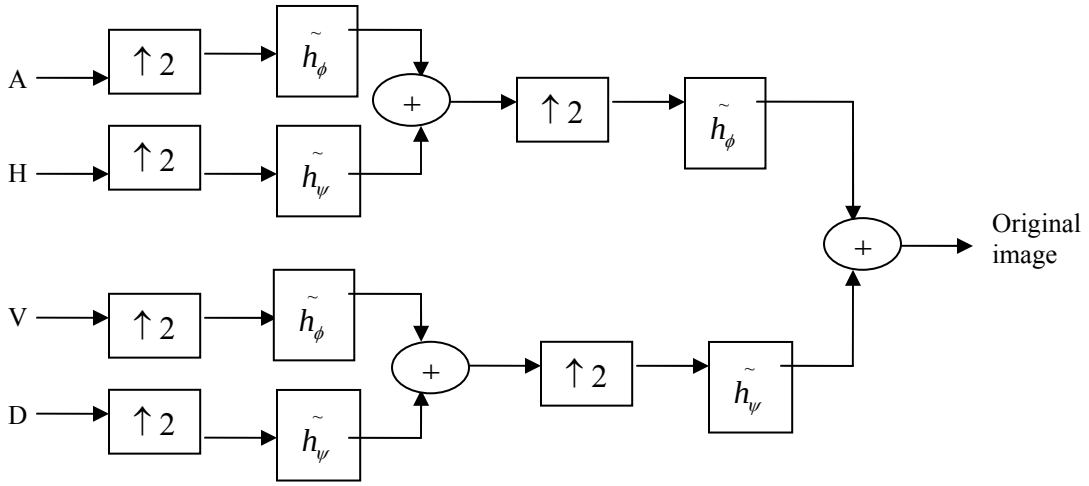


Figure B-2 The inverse 2-D DWT

It is noted that some authors [8] use a reverse notation to denote the resolution, i.e. an increasing subscript is used to denote lower resolutions. Thus, in the above discussion on DWT if the original image has a resolution J, successive DWTs will yield images at resolutions J+1, J+2, and so on. This reverse notation is used to explain the RWT, LP, and AWT in the following sections to maintain consistency with the original articles.

B.2 The Redundant Wavelet Transform

If the wavelet transform is applied to a shifted copy of the signal, the wavelet coefficients should merely be a shifted version of the coefficients that were obtained by

www.manaraa.com

applying the wavelet transform on the original signal. This property of the wavelet transform is called shift invariance [13].

The DWT described above is not shift invariant, the wavelet coefficients of the DWT change when the signal is shifted. Shift variance results from the application of subsampling in the wavelet transform. The main step in all the wavelet transforms is convolving the signal (or image) with a filter bank to obtain the approximation and the detail images. In the DWT algorithm the output of the filtering is critically subsampled, i.e. the outputs of the filter banks are decimated by a factor of two (the most common case). This subsampling causes the coefficients to change when the input is shifted. The shift variance of the DWT can sometimes be a problem in applications like pattern recognition or image fusion. For example, the use of the DWT in image fusion is known to cause artifacts in the fused images.

The problem of shift variance is overcome by oversampling or removing the subsampling step at each scale in the DWT and instead upsampling the filters at each sacle. Since the subsampling is eliminated there is redundant data in the approximation and detail images at each scale and the transform is no longer orthogonal. Thus the oversampled DWT is known as the Redundant Wavelet Transform (RWT). Suppose the original image is at resolution j=0, then the lowpass and highpass filters at each scale are upsampled by inserting 2^j-1 zeros between its non zero coefficients [13]. Thus, the filter coefficients at each successive resolution j are given as follows:

www.manaraa.com

$$h_{\phi}^{j} = h_{\phi} \uparrow 2^{j}, \tag{69}$$

$$h_{\psi}^{j} = h_{\psi} \uparrow 2^{j}, \tag{70}$$

Figure B-3 shows the diagram of the RWT. Since, the downsampling is eliminated the A, H, V, and D images are all the same size as the original image. Thus, the RWT is both computation and data intensive compared to the DWT.

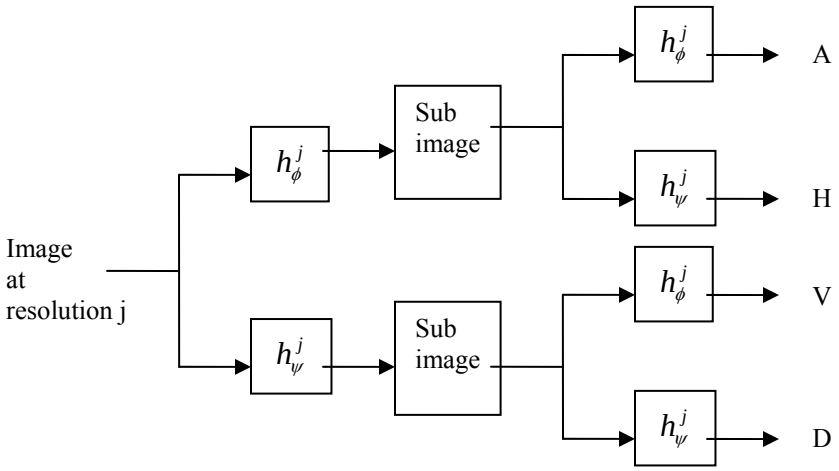


Figure B-3 The Redundant Wavelet Transform

B.3 The Laplacian Pyramid

The Laplacian Pyramid (LP) was first proposed by Burt et al. [5] for compact image representation. The basic steps of the LP are as follows:

1. Convolve the original image g_0 with a lowpass filter w (for e.g. the Gaussian filter) and subsample it by two to create a reduced lowpass version of the image – g_1,

www.manaraa.com

2. This image is then upsampled by inserting zeros in between each row and column and interpolating the missing values by convolving it with the same filter w to create the expanded lowpass image $g_1^{'}$, which is subtracted pixel by pixel from the original to give the detail image $-L_0$ given by:

$$L_0 = g_0 - g_1^{'}, \tag{71}$$

In order to achieve compression, rather than encoding g_0, the images L_0 and g_1 are encoded. Since g_1 is the lowpass version of the image it can be encoded at a reduced sampling rate and since L_0 is largely decorrelated it can be represented by far fewer bits than required to encode g_0. The above steps can be performed recursively on the lowpass and subsampled image g_1 a maximum of N number of times if the image size is 2^Nx2^N to achieve further compression. Thus the end result is a number of detail images $L_0, L_1 \cdots L_N$ and the lowpass image g_N. Each recursively obtained image in the series is smaller in size by a factor of four compared to the previous image and its center frequency reduced by an octave. Thus they can be visualized as a pyramid created from the stack of images $L_0, L_1 \cdots L_N$ one on top of another. As the detail images are a difference of two images filtered by a Gaussian or any other identical function they are called the Laplacian images and the scheme to obtain them is called the Laplacian Pyramid.

The inverse transform to obtain the original image g_0 from the N Laplacian images $L_0, L_1 \cdots L_N$ and the lowpass image g_N is as follows:

www.manaraa.com

1. g_N is upsampled by inserting zeros between the sample values and interpolating the missing values by convolving it with the filter w to obtain the image g'_N,
2. The image g'_N is added to the lowest level Laplacian image L_N to obtain the approximation image at the next upper level:

$$g_{N-1} = L_N + g'_N, \tag{72}$$

3. Steps 1 and 2 are repeated on the Laplacian images $L_0, L_1 \cdots L_{N-1}$ to obtain the original image.

B.4 The A Trous Wavelet Transform

The A Trous Wavelet Transform (AWT) is very similar to the LP except that the lowpass images are never subsampled. Thus, the approximation and detail images at consecutive scales are the same size as the original image [26]. Since the approximation and detail images of the AWT are not subsampled, in an effort to avoid confusion a different notation will be used to denote them – a_j and d_j, respectively. Prior to filtering the image at each scale j, the filter w is upsampled by inserting 2^j-1 zeros between its sample values and consequently, it is denoted by w_j to indicate this. The following two equations describe how the approximation and detail images for each scale are obtained.

www.manaraa.com

$$a_j = a_{j-1} * w_j \,, \tag{73}$$

$$d_j = a_{j-1} - a_j \,, \tag{74}$$

The inverse transform to obtain the original image a_0 is very simple. The detail images at each scale are simply added one by one to the approximation image a_N at the lowest level:

$$a_0 = a_N + d_0 + d_1 + \cdots + d_N \,, \tag{75}$$

www.manaraa.com