Image Fusion by means of DWT for Improving Classification Accuracy of RS Data

A. L. Choodarathnakara, Dr. T. Ashok Kumar, Dr. Shivaprakash Koliwad, Dr. C. G. Patil

Abstract—Fusion of Remote Sensing (RS) Images is an important process of integrating the spectral information of a single sensor or the information from different kinds of sensors. The image fusion results in a new image which is more suitable for human and machine perception or further image-processing tasks such as segmentation, feature extraction and object recognition. The fused image should preserve, as closely as possible, all relevant information contained in the input images. The fusion process should not introduce any artifacts or inconsistencies, which can be discarded or mislead the human observer. In the fused image, irrelevant features and noise should be suppressed to a maximum extent.

This paper explains how Discrete Wavelet Transform (DWT) can be used for merging the lower frequency component of a multi-spectral image and its higher spatial resolution images by means of rules. Then, using DWT fused image can be used to the spectral post processing such as classification which leads to more accurate and precise results.

Index Terms—Image Fusion, Remote Sensing, Discrete Wavelet Transform, Image Classification

I. INTRODUCTION

Image fusion is a process of combining two or more images, obtained from new and composite images using a certain algorithm. Image fusion is to integrate different data in order to obtain more information than that can be derived from each of the single sensor data alone. Image fusion has been applied to achieve a number of objective like image sharpening, improving geometric correction, complete data set for improved classification, change detection, substitute missing information, replace defective data. Data fusion is a formal framework in which are expressed means and tools for the alliance of data originating from different source gives “different quality “ means that will depend upon the application. Some generic requirement can be imposed on the fusion result.

• The fusion process should not introduce any artifacts or inconsistencies, which can be discarded or mislead the human observer.

• The fused image should preserve as closely as possible, all relevant information contained in the input images.

• In the fused image, irrelevant features and noise should be suppressed to a maximum extent.

Image fusion is a technique of obtaining images with high spatial and spectral resolution from low spatial resolution and high spatial resolution images. There is often an inverse relationship between the spectral and spatial resolution of the image. Due to the demand for higher classification accuracy and the need in enhanced positioning precision there is always a need to improve the spectral and spatial resolution of remotely sensed imagery.

II. DISCRETE WAVELET TRANSFORM

The DWT fusion methods provide computationally efficient image fusion techniques. Various fusion rules for the selection and combination of sub-band coefficients increase the quality (perceptual and quantitatively measurable) of image fusion in specific applications [1, 2].

DWT is constructed to record the selection results based on a maximum selection rule. Rather than using a binary decision, the resulting coefficients are given by a weighted average based on the local activity levels in each of the images sub-band. Another method called contrast sensitivity fusion is used. This method uses a weighted energy in the human perceptual domain, where the perceptual domain is based upon the frequency response, i.e. contrast sensitivity, of the human visual system.

Wavelets can realize by iteration of filter with rescaling. The resolution of the signal, which is a measure of the amount of detail information in the signal, is determined by the filtering operation, the scale is determined by up sampling and down sampling. Wavelet functions are normally used when calculating the Discrete Wavelet Transform.

\[
\hat{D} \left( 2^{j} - k \right) \cong \sum_{k \in Z} h(t) D(2t - k) \ldots (1)
\]

Manuscript received Nov 20, 2012.


B. Dr. T. Ashok Kumar, Department of Electronics & Communication Engineering, V. K. Institute of Technology, Mysore-570001, INDIA.

Dr. Shivaprakash Koliwad, Department of Electronics & Communication Engineering, Malnad College of Engineering, Hassan-573201, INDIA.

Dr. C. G. Patil, Master Control Facility, Hassan-573201, INDIA.
Wavelets can be determined from corresponding low-pass and high-pass filters of a perfect reconstruction filter bank. The scaling function can be calculated from the low pass filter.

\[ \Psi(t) = 2^k \sum_{n=-\infty}^{\infty} h(n) \phi(2t-n) \]

(2)

![Diagram of Discrete Wavelet Transform Decomposition Level](image)

After one level of decomposition, there will be four frequency bands, namely Low-Low (LL), Low-High (LH), High-Low (HL) and High-High (HH) as shown in Fig. 2. The next level decomposition is just applied to the LL band of the current decomposition stage, which forms a recursive decomposition procedure. Thus N-level decomposition can finally have \(3N+1\) different frequency bands, which include \(3N\) high frequency bands and just one LL frequency band.

![Wavelet Decomposition Tree](image)

The DWT is computed by successive low pass and high pass filtering of the discrete time domain signal as shown in Fig. 3. At each decomposition level, the half band filters produce signal spanning only half the frequency band. This doubles the frequency resolution as the uncertainty in frequency is reduced by half. In according with Nyquist rule the original signal has highest frequency of \(w\) which requires a sampling frequency of \(2w\) radians, then it now has a highest frequency of \(w/2\) radians. It can be now be sampled at a frequency of \(w\) radians thus discarding half the sample with no loss of information. This decimation by 2 halves the time resolution as the entire signal is now represented by only half the number of samples. Thus while the half band low pass filtering removes half of the frequencies and thus halves the resolution, the decimation by two doubles the scale. The filtering and decimation process is continued until the desired level is reached. The maximum number of levels depends on the length of the signal. The DWT of the original signal is the obtained by concatenating all the coefficients \(a[n]\) and \(d[n]\) starting from the last level of decomposition and for inverse reverse operation can be done [7].

### III. STUDY AREA AND DATA PRODUCTS

#### A. Study Area

Because the study area considered for our work is a semi-urban area of Arsikere situated in Hassan District, Karnataka State, India and its geographical coordinates are 13° 18’ 50” North, 76° 15’ 22” East and its original name (with diacritics) is Arsikere. It has an average elevation of 807 meters (2647 feet). The image dimension of the study area is 607×645 pixels in MS data and a visual band aerial photograph of the study area is 726×607 pixels of PAN data.

![Visual Band Aerial Photograph of Arsikere Study Area](image)

#### B. Data Products

Table 1 gives the specification of image data products used in this project work. The data are of LISS-IV (Linear Imaging and Self Scanning) sensor of IRS P-6 (Indian Remote Sensing Satellite). The satellite data were obtained from the Master Control Facility, Hassan, India. Another data used is the Visual band multi spectral data downloaded from www.wikimapia.com.

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Satellite and Data type</th>
<th>Date of Acquisition</th>
<th>Spectral Resolution</th>
<th>Spatial Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>IRS P-6 (Resourcesat1) Multi-spectral</td>
<td>July 2002</td>
<td>Green(0.52-0.59µm); Red (0.62-0.68 µm); Infrared(0.77-0.86 µm)</td>
<td>5.8m</td>
</tr>
<tr>
<td>2.</td>
<td>Wikimapia.com Panchromatic</td>
<td>March 2010</td>
<td>-----</td>
<td>2.5m</td>
</tr>
</tbody>
</table>

TABLE I

**Details of the Data Products Used**
FUSION RULES AND ALGORITHM

Fusion rule is defined as the process of image fusion, while fusion operates is fusion method which is given in the fusion rule. Fusion rule can be computed by using activity rule. This is defined below

\[ A_f(p) = f(U(p)) \]  \( \ldots \) \( (3) \)

Where, \( P \) is the position of the image decomposed coefficient, \( U(p) \) is a neighborhood which is centered at point \( p \) and \( f \) is a function of computing activity level.

There are three methods of computing the activity level of a coefficient: Coefficient based, window based and region based measures. Generally, an image has its representation denoted as \( ID \). Let \( P = (m, n, j, l) \) indicate the index corresponding to a particular coefficient, \( m \) and \( n \) indicate the spatial position in a given frequency band, \( j \) is the decomposition level and \( l \) is the frequency band of the representation.

Some kind of fusion rules are used to fuse the multiscale and multidirectional images from decomposition. The low frequency part includes most of the background information; the high frequency parts include most of the image detail information. The correlation between a pixel is about the region in which the pixel and its neighboring pixel is often higher than others, so the fusion rules is about the region in which the pixel is the neighboring pixels.

A. The Fusion Rule for Low Frequency Area

In the process of image fusion the maximum wavelet coefficient with low frequency. This text makes use of spatial frequency and contrast degree to make sure the low frequency of image fusion. Spatial frequency reflects the total active degree of the image spatial frequency reflects the total active degree of the image. Suppose an image size be \( M \times N \), the worth of the image for the ash degree of place is \( (m, n) \).

Spatial frequency is calculated by:

\[ SF = \sqrt{SF^2 + CF^2} \]  \( \ldots \) \( (4) \)

The RF is row frequency and the CF is the list frequency:

\[ RF = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=2}^{N} F(m, n - 1) \]  \( \ldots \) \( (5) \)

\[ CF = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=2}^{N} F(m, n - 1) \]  \( \ldots \) \( (6) \)

In the formula, \( I \) is the image bright degree, \( I_s \) means the low frequency part of the image bright degree and \( I_h \) means the high frequency part of the image bright degree.

The source image \( A \) and \( B \) will be decomposed with the Mallat algorithm. Assume the low frequency is \( A_L \) and \( B_L \), they are made into the sub-images whose size is \( M \times N \) respectively, the subimage are recorded respectively for the \( A_{Lx} \) and \( B_{Lx} \).

Calculate the spatial frequency of \( SF_{ak} \) and \( SF_{bk} \) and the contrast \( C_{ak} \) and \( C_{bk} \) with \( A_{Lx} \) and \( B_{Lx} \) respectively. The No. \( k \) sub image is determined by the formula

\[ C = (I-I_B)/I_B = I_H/I_B \]  \( \ldots \) \( (7) \)

In the formula, \( I_B \) means image bright degree, means the low frequency part of the image bright degree and \( I_H \) means the high frequency part of the image bright degree.

B. The Fusion Rule for High Frequency Area

The absolute maximum of the wavelet coefficient should be matching to the maximal contrast of the image edge characteristic that is the important information of the image. Therefore we choose the maximal coefficient wavelet rule to carry on the fusion in the high frequency area for the image \( A \) we define the characteristic variable \( S \) of the wavelet coefficient.

\[ S_f^j(q, K) = \max_{m,n} \left\{ S_f^j(q, p) \right\} \]  \( \ldots \) \( (8) \)

In the formula, \( j \) represents the number of the layers of the wavelet coefficient, \( K = (m, n) \) is the spatial position of the wavelet coefficient. \( Q \) is a small area by the center \( K \). The same way defines image \( B \). In the process of the image fusion, we select the wavelet coefficient with the maximal characteristic variable. By this we can get these formulas;

\[ Z_f^j(A, K) = \left\{ \begin{array}{ll}
0, & S_f^j(A, K) \leq S_f^j(B, K) \\
1, & S_f^j(A, K) > S_f^j(B, K)
\end{array} \right. \]  \( \ldots \) \( (9) \)

\[ Z_f^j(B, K) = \left\{ \begin{array}{ll}
0, & S_f^j(B, K) \leq S_f^j(A, K) \\
1, & S_f^j(B, K) > S_f^j(A, K)
\end{array} \right. \]  \( \ldots \) \( (10) \)

According to the majority representation method, the above two formula are modified

\[ Z_f^j(A, K) = \left\{ \begin{array}{ll}
1, & Z_f^j(A, K) \geq 5 \\
0, & Z_f^j(A, K) < 5
\end{array} \right. \]  \( \ldots \) \( (11) \)

\[ Z_f^j(B, K) = \left\{ \begin{array}{ll}
1, & Z_f^j(B, K) = 0 \\
0, & Z_f^j(B, K) = 1
\end{array} \right. \]  \( \ldots \) \( (12) \)

Referring to the above formulas, the fusion rule to the high frequency is founded as;

\[ D(K) = Z_f^j(A, K) \cdot Z_f^j(A, K) + Z_f^j(B, K) \cdot Z_f^j(B, K) \]  \( \ldots \) \( (13) \)
C. Panchromatic & Multi-spectral (MS) Image Fusion

Remote sensing images can be acquired in two different modes: either the panchromatic (P) mode with high spatial resolutions (HR) of 10 m pixel (SPOT), 2.5 m pixel (IRS) or even 1 m pixel (IKONOS), either the MS mode with much lower spatial resolution (LR) of 30 m pixel (LANDSAT TM), 5.8 m pixel (IRS), 20 m pixel (SPOT) or 4 m pixel (IKONOS). The P images are characterized by very high spatial information content well suited for intermediate scale mapping applications and urban analysis. The MS images provide the essential spectral information for smaller scale thematic mapping applications such as land use surveys. In order to take benefit of the high spatial information content of the P images and the essential spectral information of lower resolution MS images, fusion of these two types of images can be performed in order to produce pseudo-HRMS (High Resolution Multi spectral Image) images.

- Match the histogram of P according to the three MS bands to obtain the corresponding PSMi at 2.5 m pixel of spatial resolution.
- Apply DWT twice for decomposition levels to each PSMi to obtain the approximation PSAi2 coefficients at the second resolution level and the PSDij details coefficients i = 1, 2, 3 (bands), j = h, v, d (horizontal, vertical and diagonal, respectively) and k = 1, 2 (decomposition level). The spatial resolution levels are 2.5 m pixel and 5.8 m pixel for k = 1, 2, respectively.
- Select and merge the corresponding PSAi2 with MSi by WA by taking a zero weight value for PSAi2 and the unitary value for MSi to obtain the fused scale PSAF. The WA carried out is a mere substitution of PSAi2 with MSi, i.e. in this particular example MSi = PSAF.

D. Fusion Algorithm

Step A:
Co-registers both images and resamples the according low resolution image to make its pixel size equal to that of the high resolution image, in order to get perfectly superposable images.

Step B:

Stage I: Read the two source images A and B to be fused. Perform independent wavelet decomposition of the two images until level L to get approximation (LLl) and detail (LHl, HLl, HHl) coefficients for l=1,2,….L.

Stage II: Select pixel based algorithm for approximations (LLl) which involves fusion based on taking the maximum valued pixels from approximations of source images A and B.

\[
LL_f^l = \max\{LL_A^l(i,j), LL_B^l(i,j)\} \quad \ldots \quad (14)
\]

Here is the fused and and are the input approximations, i and j represents the pixel position of the subimages. LL LA LL LB LL.

A binary decision map is formulated based on maximum valued pixels between the approximations. The decision rule Df for fusion of approximation coefficient in the two source images A and B is thus given by:

\[
D_f(i,j) = \begin{cases} 
1, & \text{if } d_f(i,j) = 0, \\
0, & \text{otherwise} 
\end{cases} \quad \ldots \quad (15)
\]

A small window of size is selected from the detail subband based on whether the type of filter mask used is square or rectangular.

Perform region level fusion of details by applying square and averaging filter mask to detail coefficients. The resultant coefficients are added from each subband.

\[
\begin{align*}
LL_f^l & = \text{mask}(LL_A^{l+1}) + \text{mask}(LL_B^{l+1}) \\
\text{HL}_f^l & = \text{mask}(HL_A^{l+1}) + \text{mask}(HL_B^{l+1}) \\
\text{HH}_f^l & = \text{mask}(HH_A^{l+1}) + \text{mask}(HH_B^{l+1}) 
\end{align*} \quad \ldots \quad (16)
\]

Here, LHl, LHlA, LHlB are vertical frequencies HLl, HLlA, HLlB are horizontal high frequencies HHl, HHlA, HHlB are diagonal high frequencies of the fused and input detail subband respectively.

Stage III: We obtain the final fused transform corresponding to approximations through pixel rules and the vertical, horizontal and diagonal details LHlA, HLlB and HHl by mask based fusion were l = 1, 2, ……L.

The new coefficient matrix is obtained by concatenating fused approximation details.

Fused image is reconstructed using inverse discrete wavelet transform and displayed.

Step B:
Insert the spatial information of the high resolution image into the lower resolution image one through the inverse wavelet transform.

V. RESULTS AND DISCUSSION

The fusion effect method is relatively ideal, its transformation restructuring is actually a process of the high pass and the low pass filter, and, to a certain extent, loses some edge information in the primitive image, as a result of which the ringing effect in fusion image appears. Therefore, in this method using discrete wavelet transformation is used to improve the features of human vision system. This method can not only enhance the spatial detail ability of multi-spectrum image, maintenance the spectrum information and also the integration of color and spatial features as natural.

A. Human Eye Visual System

The image is to be looked at by humans; therefore the imagery processing should follow the features of human vision system. The vision is the result that images are left in human eyes. The complexity of visual processing has not have been understood and mastered by human beings at present. But people have discovered some visual phenomena, such as visual threshold, visual masking effect. If these features of human vision system are fully used in the image processing, the image quality will be greatly improved.
B. Fusion Steps based on the features of human vision system using integration method

As for combining the discrete wavelet transformation and HIS with the feature of human vision system, a new fusion method is brought up.

- Perform the geometry correction to the low resolution image and the high resolution image respectively, use the geometry matching method based on the region, and then match the multi-spectrum image with the high resolution image.
- Perform a discreet wavelet resolving to the matched low resolution image and the high resolution image to obtain respectively the low frequency component and the high frequency detail component [4, 5].
- Perform the uniformity method fusion to the two corresponding images of the high frequency and the low frequency sub-graph respectively, the processes are:
  - Dissect the two images into some sizes for NxN blocks. Suppose A_i and B_i expresses the block of image A and B separately.
  - Calculate the uniformity measure of each block according to the formula which provides. Suppose J(A_i) and J(B_i) are respectively the uniformity measures of A_i and B_i.
  - Compare the uniformity measure of the corresponding block of the two images, and obtain the fusion image i_th block F_i. Among them, T_H is the threshold value parameter.
  - In turn carry on the above operation to all image blocks, the new fusion image is obtained [6].
- Carry on the inversion to the obtained high frequency and the low frequency sub-graph, and obtain the fused multi-spectrum image.

C. Visual Analysis of the Fusion Result

Fig. 6 Snapshot of test input image of Multispectral image with 5.8m resolution with size 741*596.

The resulting images obtained from Laplacian Pyramid approach had the erosion scars and debris deposition areas enhanced in relation to the other targets in the scene. These features presented light and pale tones, contrasting with forest, natural fields, etc. However, for both approaches one can verify that it was not possible to obtain a good detail in the landslide areas as shown in Fig. 7. For example, grass, natural fields and vegetation cover in the early growing stages were not well discriminated. It was also possible to notice a defocusing in the image, which made the erosion scars look like stains instead of linear features.

On the other hand, the DWT image evidenced the erosion scars, showing them with clear and pale tones as well as with rectilinear and elongated shapes (Fig. 8). The boundaries of the scar areas became distinct from the boundaries of the other targets. Due to the preservation of spectral characteristics, it was possible to discriminate more precisely other targets existing in the scene. After the spatial resolution enhancement, it was also possible to identify several small features of landslides that were not identified using other fusion approaches. For example the colour of the wavelet fused result is close to that of original colour image in vegetation areas, but with obviously weak colour intensity in residence areas. The railway coaches, roads and building roofs have almost the same colour except the brightness variation. Trees in residence areas can be hardly recognized.
The spatial quality of the fusion results has also been analyzed by enlarging the fused images and the panchromatic images. In all of the fusion results, Railway coach, roads, building corners and other sharp edges can be seen as clear as in the original panchromatic images. This indicates the spatial qualities of all the fusion techniques being tested are similar, or the same.

It is important to clarify that this method used for testing fusion performance is only representative for visual applications where the source images are in focus in different areas. And the approach based in the Discreet Wavelet Transform was developed specifically for this kind of applications, so it is understandable that its results were slightly better than the other approaches. However, opposite to the others, this technique was found definitely unreliable for other image fusion applications where the source images come from different sensors. The Fig. 9 and Table 3 gives the RMSE values for the two different fusion techniques with number of scales.

VI. CONCLUSION

Most of the image fusion approaches were found reliable fusion methods in conjunction they gave acceptable results in multi-sensor fusion schemes and it is difficult to conclude which method is the best one for a certain application. For this purpose some visual tests were carried out, where a group of individuals express their subjective preferences between couples of images obtained with different fusion methods.

In DWT the different parts of the image after decomposition, different fusion rules are adopted. The intensity of the high frequency information is merged with the low frequency. In DWT the different parts of the image after decomposition, different fusion rules are adopted. The intensity of the high frequency information is merged with the low frequency information to match the intensity component to get the feasible and effective method for image fusion.

Since fused images are used to enhance visual information for human users, performance assessment of image fusion should be first judged by the users, based on the mission of specific applications. In situations where it is hard to take decision about the quality of fused image objective measures are calculated. Quantitative measures should only serve as a useful tool to assist human users to make difficult judgments whenever necessary. The performance assessment of image fusion should continue to be shared between qualitative and quantitative methods, with increasing weight being placed with quantitative assessment techniques.

ACKNOWLEDGMENT

It is a pleasure to recognize the many individual who have helped me in completing this paper. I sincerely express heartiest thanks to my guide Dr. Shivaprapaksh Koliwad for all the technical guidance, encouragement and support throughout this process. Also, I thank to Dr. C. G. Patil (Director of MCF) for his external support and guidance.

REFERENCES


A. L. Choodarathnakara was born in Hassan District of Karnataka State in INDIA. He received the B.E. degree in Electronics & Communication Engineering in the year 2002 and M.Tech. degree in Digital Electronics & Communication Systems in the year 2008, both from the Malnad College of Engineering, Hassan affiliated to Visvesvaraya Technological University, Belgaum, INDIA.

He is an Associate Professor of Electronics & Communication Engineering Department at Government Engineering College, Kulashtnagar of South Kodagu District in Karnataka. His subjects include Natural Language Processing, Language Learning Technologies, Text, Speech, Image, Signal & Remote Sensed Data Processing.

Asso. Prof. Choodarathnakara A. L. is a Life Member of the Indian Society of Technical Education (ISTE-43168), Indian Society of Remote Sensing (ISRS-3533) and Indian Society of Geomatics (ISG-1258).

Dr. Shiva Prakash Koliwad received the B.E. degree in 1974 from SJC, Mysore (Mysore University), M. Tech. from IIT, Delhi, and the Ph. D. degree in 2000 from Anna University, Chennai, India. His subjects of interest are Image Processing, Signal & Remote Sensed Data Processing & Bio-medical Engineering. His specialization is in Split Band Bio-feedback Technique for bio-medical signal analysis.

At present, he is a Professor in the Department of Electronics and Communication Engineering at Malnad College of Engineering, Hassan. He is in teaching since 1975, and is a life member of the Indian Society for Technical Education (ISTE).

Dr. C. G. Patil received the M.Sc. degree in Physics from IIT, Kanpur and the Ph. D. degree from Karnataka University, Dharwad, India. His areas of interest are Satellite Image Processing, Remote Sensed Data Processing, Spacecraft Altitude Control, Communication, Mission Management and Expert Systems. His specialization is in Transport Properties in Ferromagnetic Semiconductors.

He is serving Master Control Facility since 1980 and presently is a Director of MCF, Hassan. Dr. C. G. Patil is involved in all the 17 mission operations of geostationary satellites controlled from MCF & operation director for INSAT-2A & INSAT-2B satellites. He is a recipient of Astronautical Society of India Award for the year 2001-02.

Dr. T. Ashok Kumar received the B.E. degree in Electronics and Communication Engineering from Malnad College of Engineering, Hassan (Mysore University) in the year 1990, and M.E. in Digital Electronics form SDM CET, Dharwad (Karnataka University), India, in 1995.

He is a Professor of Electronics and Communication Engineering, and currently he is serving as the Principal of Vivekananda College of Engineering and Technology, Puttur, Karnataka. He is in teaching for 20 years. His subjects of interest include Image Processing, Communication Engineering, Data Mining and Remote Sensing.

Prof. T. Ashok Kumar is a life member of Indian Society of Geomatics (ISG), Indian Society of Remote Sensing (ISRS), ISTE, and member of IEEE.