Remote Sensing Image Fusion for Different Spectral and Spatial Resolutions with Bilinear Resampling Wavelet Transform

Zhou Qianxiang  Jing Zhongliang
Institute of Aerospace Information and Control, School of Electronics and Information Technology, Shanghai Jiao Tong University, Shanghai, CHN, 200030,
zqxg@sjtu.edu.cn, zljing@mail.sjtu.edu.cn

Jiang Shizhong
Institute of Space Medico-Engineering, Beijing, CHN, 100094

Abstract: It is an important way that some remote sensing images of different spatial and spectral resolutions are fused to satisfy the requirement of general application. In order to achieve a good fusion result, low spatial spectral images should be resampled. At present, nearest neighbor resampling is often adopted which has some effects on the precision of new image. In this paper, an image fusion method is proposed with bilinear resampling wavelet (BRW) transform, and compared with nearest neighbor resampling wavelet transform, IHS transform and Brovery transform. On the platform of ENVI/IDL simulations show that the BRW method is of good performance of preserving the spectral and spatial resolutions for remote sensing images, with lowest the lost of spectral information.

Keywords: Remote Sensing Image, Wavelet Transform, Fusion Algorithm, Bilinear Resampling.

1 Introduction

In the recent twenty years, remote sensing technology has achieved great progresses, which means that variable airplane born or satellite born remote sensing systems have been developed with high spectral or spatial resolutions for different application objectives. The field of remote sensing is a continuously market with applications like vegetation mapping and observation. However, as a result of the demand for higher classification accuracy and the need in enhanced positioning precision (e.g. for geoscience information systems), there is always a need to improve the spectral and spatial resolutions of remotely sensed imagery. There are two ways of realizing this objective. One way is application of high spatial and spectral resolution of remote sensors. For example, multispectral France SPOT scene consists of three bands called XS1, XS2 and XS3. The first two bands characterize the ground cover in the visible wavelength range and the third band near infrared wavelength. These bands are acquired with a spatial resolution 20X20m per pixel, the wavelength range of the panchromatic image P10 covers the bands XS1 and XS3 with a ground resolution of 10X10m[1, 2]. Indian remote sensing satellite (IRS) C, D Pan sensor sacrifices swath width for its higher resolution, both of which produce 5.8-meter panchromatic (0.5 or 0.75 µm - black and white) imagery, which is resampled to five-meter pixel detail. However, IRS-LISS III can provide high spectral imagery of 23.5X23.5m resolution[3, 4]. In addition, USA IKONOS remote sensing satellite’s ground resolution of each band is 1-meter panchromatic (nominal at <26 deg off nadir) and 4-meter multispectral (nominal at <26 deg off nadir), their band wavelength differs from 0.45 to 0.90 microns. Another way of acquiring high quality of remote sensing imagery is called enhancement of the spatial resolution of multispectral image or...
multispectral image sharpening, if features of remote sensors are defined. As a result, based on multi-sensors fusion theory, multisource remote sensing image fusion has been studied from 1980’s and is becoming an important way of remote sensing image processing.

The study of remote sensing image fusion has made great progresses, including variable idiographic algorithms and widely used evaluation indexes for fused image. After having summarized the concept, methodology and application of image fusion in detail, prof. Pohl and Van G think that image fusion can be divided into three levels: pixel, feature and decision level fusion. The indexes of image evaluation may be statistical comparison, visual comparison, graphical comparison, least mean square deviation computation, image entropy, image noise index, etc [5, 6]. From results of literature searching in IEEE, EI and international journal of remote sensing, there are a great many of remote sensing techniques to some practical applications. Some of them include intensity-hue-saturation (IHS) transform fusion, principal component analysis (PCA) fusion, Brovery transform fusion, wavelet transform fusion, etc., the basis of which is that before image fusion, all images to be processed must be registered and georeferenced and low spatial resolution multispectral images should be resampled into new images with the same resolution as Pan images [4, 7]. In general, the resampled method is nearest neighbor. As it being a simple one to some extent, compared with other resampled ways, the calculation precision of this method is the lowest and the corresponding loss of image spectral information is the highest. But nearest neighbor resampling computing speed is the fastest. Thereby, two factors should be taken into consideration when the remote sensing images are fused: one is the loss of image spectrum, the other is the effect of image resampling method on computer speed.

Wavelet transform is a global mathematic analysis in that it has a good position capability in both time and frequency domains. If the reduced length is gradually introduced when detailed images of wavelet transform are decomposed, any detail of the image to be processed can be focused. As a result, it is widely put into use in remote sensing image processing. In this paper, results of simulations on the platform of ENVI/IDL are reported as follows: Firstly, a Land TM multispectral image and a SPOT Pan image are fused with bilinear resampling wavelet transform. Secondly, this method is compared with other fusion ones such as wavelet transform with nearest neighbor resampling, IHS transform and Brovery transform according to the image evaluation index. In all, it can be concluded that the bilinear resampling wavelet transform fusion is of the best fusion result.

2 Fusion Algorithm

Normally, the approach of remote sensing images fusion includes the following two steps: Firstly, low spatial multispectral images are resampled into new images with the same resolution as Pan images. Secondly, these images are fused with a fusion algorithm (wavelet transform, PCA, IHS transform, etc.). Because the remote sensing images are very large, nearest neighbor resampling method is always chosen first and foremost considering computing speed. However, being only simply replacement of image pixel, nearest neighbor resampling will affects the fused image quality [6, 7, 8, 9]. Therefore, a new image resampling method should be applied. In this paper, bilinear resampling is studied, which can be shown as figure 1. According to classical mathematic theory, it is expressed as follows:

For image pixel \((x, y)\) and its nearby pixels \((0,0), (0,1), (1,0)\) and \((1,1)\), their gray values are \(f(x, y), f(0,0), f(0,1), f(1,0)\) and
respectively. They are calculated as follows:

\[ f(x, y) = f(x, 0) + x[f(1, 0) - f(0, 0)] \]  (1)

Similarly, \( f(x, 1) \) can also be got with the same way.

\[ f(x, 1) = f(0, 1) + x[f(1, 1) - f(0, 1)] \]  (2)

Finally, \( f(x, y) \) is got with vertical linear sampling.

\[ f(x, y) = f(x, 0) + x[f(x, 1) - f(x, 0)] \]  (3)

After the low spatial multispectral images having been resampled, all images have the same high spatial resolution. They are fused with wavelet transform, which is shown as figure 2.

In figure 2, an explicit expression of wavelet transform fusion framework has been given, where all images are decomposed into two parts: approximate and detailed coefficient wavelet basis. The second component includes horizontal, vertical and diagonal coefficients. The essence of image fusion is that to some extent the wavelet coefficients of low spatial resolution multispectral images are merged with that of high spatial resolution image. Then the fused image’s wavelet coefficients are achieved. If they are reconstructed, remote sensing image fusion will be finished.

At present, using detailed images \( (C_h, C_v, \)
and $C_d$), the detailed images that are going to be replaced in the multispectral decomposition have been generated with three methods\[10,11\].

**Method I:**
$$C_h = C^p_h, C_v = C^p_v \text{ and } C_d = C^p_d.$$  

**Methods II:**
$$C_h = a_h C^p_h + b_h \quad \text{ and } \quad C_v = a_v C^p_v + b_v$$
$$C_d = a_d C^p_d + b_d.$$  

where $a_h, b_h$ are computed using $C^p_h$ and $C^M_v$.

**Method III:**
$$C_h = \frac{C^p_h + C^M_h}{2}, \quad C_v = \frac{C^p_v + C^M_v}{2}$$
$$\text{and } \quad C_d = \frac{C^p_d + C^M_d}{2}.$$  

The remote sensing image decomposing and reconstructing are computed according to the wavelet transform theory, which are expressed as follows:

Let $\psi(t) \in L^2(R) \quad (L^2(R) \text{ is the square integrable real space}),$ the FFT of which is shown as $\psi^*(\omega)$, if $\psi^*(\omega)$ satisfies:

$$C_v = \int_{\mathbb{R}} \frac{|\psi^*(\omega)|^2}{|\omega|} d\omega < \infty \quad (4)$$

$\psi(t)$ is called a basic wavelet or a mother wavelet. After it have been scaled and shifted, a series of wavelet $\psi_{a,b}(t)$ will be got:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi(t-b) \quad a, b \in \mathbb{R} ; \quad a \neq 0 \quad (5)$$

where $a$ is called scale factor and $b$ being shifting one.

For any function $f(t) \in L^2(R)$, its continuous wavelet transform $W_f(a, b)$ will be as follows.

$$W_f(a, b) = \left| a \right|^{-\frac{1}{2}} \int_{\mathbb{R}} f(t) \psi^\ast \left( \frac{t-b}{a} \right) dt \quad (6)$$

where sign "-" means complex conjugate.

There are many applicable wavelet families, the representatives of which are Harr, Daubechies, Morlet, Symlets A, Mexican Hat, Meyer and Coiflet. They can be chosen according to the requirement of application.

As expressed above, when images are fused, the most important work is to reconstruct new fused image from coefficient wavelet basis\[12\]. In general, an image can be regarded as a discrete two-dimensional function $f(x, y)$, which is also shown in figure 2. It can be concluded that the low frequency of scale j is divided into four parts: the low frequency of scale j+1 and high frequency of three directions (horizontal, vertical and diagonal). Mallat fast computing algorithm is often used for decomposition of a two-dimensional function, which means that function $f$ is decomposed according to different frequency channels and each frequency is decomposed by its phase. The higher the frequency is, the more detailedly its phase is divided, vice versa. Compared with signal processing, those divisions can be regarded that images are low-pass and high-pass filtered in horizontal direction, then they are down sampled. Finally, they are filtered with the same way in vertical direction. The formulas are expressed as follows.

Let $f(x, y) \in V_j^2 (j \in \mathbb{Z})$, its projection in $V_j^2$ space is called $A_j f(x, y)$, where $V_j^2$ is a multiscale space division of $L^2(R)$, then

$$A_j f(x, y) = A_{j+1} f(x, y) + D_{j+1}^1 f(x, y) + D_{j+1}^2 f(x, y) + D_{j+1}^3 f(x, y) \quad (7)$$

where
of ENVI/IDL

ENVI/IDL software is a synthetic platform developed by Research Systems, Inc. USA mainly for remote sensing image processing. It is capable of general analysis of remote sensing images and provides special API functions for users. In this paper, a SPOT Pan image and Land TM multispectral image are studied. The file size of SPOT image is 2820X1569 bytes, its resolution being 10 m. Bands 4, 3, 2 of Land TM image are used, the file size of which is 1007X560 bytes and its resolution being 28 m. They are shown in figure 3(a) and 3(b) respectively. After having been registered and georeferenced, these two images are fused with nearest neighbor and bilinear resampling Symlet 4 wavelet(figure (c) and 3(d) respectively), IHS and Brovey transform for comparison. All fused images are evaluated with the following indexes, results of which are shown in table 1.

(1) Image entropy

Claude Shannon was the first person to introduce entropy in the quantification of information. Shannon(1948) employed the probabilistic concept in modeling message communication. He believed that a particular message is one element from a set of all possible messages\(^9\). The main purpose of image fusion is to improve image information. Therefore, Shannon’s entropy can be put into use when the fused image is evaluated, which is defined as:

\[
I = -\sum_{i=0}^{255} P(i) \log P(i) \tag{9}
\]

where \( P(i) \) means probability of pixel value \( i \) in the image.

3 Simulation results on platform

\[
A_{j+1} = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} C_{j+1}(m_1, m_2) \psi_{j+1}(m_1, m_2)
\]

\[
D_{j+1} = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} D_{j+1}(m_1, m_2) \psi_{j+1}(m_1, m_2)
\]

\( i = 1, 2, 3 \)

\( C_{j+1}(m_1, m_2) \) means approximate component of image decomposition with binary grid \( 2^{j+1} \) scale.

\( D_{j+1}(m_1, m_2) \) means detailed component of image decomposition with binary grid \( 2^{j+1} \) scale. If \( i \) varies from 1 to 3, \( D_{j+1}(m_1, m_2) \) is called as horizontal, vertical and diagonal wavelet coefficient respectively. 

\( (m_1, m_2) \) means the position of image pixel.

If \( H_r, H_c \) and \( G_r, G_c \) are called mirror conjugate filters, \( H \) and \( G \) only affect rows or columns of array

\[
\{C_j(m_1, m_2), (m_1, m_2) \in \mathbb{Z}^2\}
\]

the binary Mallat decomposition will be shown as:

\[
\begin{align*}
C_{j+1} &= H_r H_c C_j \\
D_{j+1}^1 &= H_r G_c C_j \\
D_{j+1}^2 &= G_r H_c C_j \\
D_{j+1}^3 &= G_r G_c C_j
\end{align*} \tag{8}
\]

Correspondingly, its reconstruction is expressed as:

\[
C_j = \overline{H_r} \overline{H_c} C_{j+1} + \overline{H_r} \overline{G_c} D_{j+1}^1 + \overline{G_r} \overline{H_c} D_{j+1}^2 + \overline{G_r} \overline{G_c} D_{j+1}^3
\]

where \( \overline{H}, \overline{G} \) are conjugate of \( H, G \) respectively.

3 Simulation results on platform
(2) Image Clarity

Image clarity means its mean grads, which can reflect image tiny detailed difference and texture change. It is expressed as:

$$\nabla G = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left( \frac{\partial f(i, j)}{\partial x} \right)^2 + \left( \frac{\partial f(i, j)}{\partial y} \right)^2 \right)^{1/2}$$

Where \( \frac{\partial f(i, j)}{\partial x} \) and \( \frac{\partial f(i, j)}{\partial y} \) are one-order differential of pixel \((i, j)\) in x and y direction respectively.

(3) Correlation coefficient \( \rho_{xy} \) between fused image and original images

Let each gray value of pixel in the fused image be random variable \( X \) and the corresponding pixel value be random variable \( Y \). Their expectation and covariance are \( E_X, E_Y \) and \( D_X, D_Y \) respectively. If expectation of random variable \((X, Y)\) is set to be \( E(XY)\), correlation coefficient \( \rho_{xy} \) is defined as:

$$\rho_{xy} = \frac{E(XY) - E_X E_Y}{D_X D_Y}$$

In order to improve the computing speed, the two images are sliced into ten pieces to be loaded into the computer for fusion one by one. With this way, the computing capacity of computer for algorithm complexity will be traded off with the image size.

From Table 1, it can be concluded that four fusion approaches – nearest neighbor resampling wavelet, bilinear resampling wavelet, IHS and Brovery are used to create fusion image, each technique results in fused image with varied characters. It is believed that the correlation coefficient directly indicates the amount of spectral content preserved. As image spectral information is concerned, we can infer that with bilinear resampling wavelet about 85% of the Land TM image seems to be preserved in fused image. In other words, about 15% spectral content is lost on fusion using bilinear resampling wavelet transform method. In a similar fashion, Brovery transform results in about 51% spectral content being preserved in fusion. Thereby, if the lost spectral content of image is ordered from high to low loss, related fusion methods will be Brovery transform → IHS → nearest neighbor resampling wavelet → bilinear resampling wavelet. In addition, whatever kind of wavelet fusion approach is chosen, the spectral content loss resulted by both
of them are less than the other two fusion ways. This means that wavelet transform fusion can preserve spectral content of remote sensing image effectively with improving its spatial resolution. The same result can also be achieved by indexes of image clarity and entropy.

### Table 1 Evaluation of Fused Images

<table>
<thead>
<tr>
<th>Fusion approaches</th>
<th>Correlation Coefficient Between Fused and Original Images</th>
<th>Image Clarity</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Land TM</td>
<td>SPOT</td>
<td>Fused Image</td>
</tr>
<tr>
<td></td>
<td>Band 4</td>
<td>Band 3</td>
<td>Band 2</td>
</tr>
<tr>
<td>Nearest neighbor Wavelet</td>
<td>0.79</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>Bilinear Wavelet</td>
<td>0.83</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>HIS</td>
<td>0.67</td>
<td>0.65</td>
<td>0.66</td>
</tr>
<tr>
<td>BROVERY</td>
<td>0.51</td>
<td>0.53</td>
<td>0.55</td>
</tr>
</tbody>
</table>

### 4 Conclusions

As Pan image and multispectral remote sensing image have different features, if they are fused to create a new image, its detail and texture will be shown more explicitly. Thereby, it can be concluded that multisource remote sensing image fusion is an important way to integrate every kind of sensors information.

The main factor which constraints the multispectral image resampling algorithm application may be computing speed of computer. Because the file size of remote sensing image is always hundreds of M-bytes, it is important to optimize the fusion computation. In this paper, the way with which that the multispectral image is sliced and loaded into computer one by one can resolve slow computing speed when multispectral image is reampled with bilinear sampling.

Wavelet analysis is a time/frequency tool with feature of multi-channel resolution. From the fused image quality, it is specially suitable for image fusion. At present, it is generally applied into pixel level fusion. For feature and decision level fusion, it is necessary to study its application further.

### References


