# Image fusion using non-separable wavelet frame

**Hong Wang (王 宏), Zhongliang Jing (敬忠良), and Jianxun Li (李建勋)**

*Institute of Aerospace Information and Control, Shanghai Jiaotong University, Shanghai 200030*

Received March 17, 2003

In this paper, an image fusion method is proposed based on the non-separable wavelet frame (NWF) for merging a high-resolution panchromatic image and a low-resolution multispectral image. The low-frequency part of the panchromatic image is directly substituted by multispectral image. As a result, the multispectral information of the multispectral image can be preserved effectively in the fused image. Due to multiscale method for enhancing the high-frequency parts of the panchromatic image, spatial information of the fused image can be improved. Experimental results indicate that the proposed method outperforms the intensity-hue-saturation (IHS) transform, discrete wavelet transform and separable wavelet frame in preserving spectral and spatial information.

*OCIS codes:* 100.2980, 100.7410, 350.2660.

Different image sensors provide different data with different spectral and spatial resolution. Multispectral imaging sensors collect poor spatial resolution multispetral data, while panchromatic imaging ones provide adequate spatial resolution panchromatic data. Thus, image fusion techniques have been proposed in order to allow integration of different information sources and the fused image may provide increased interpretation capabilities and more reliable results. Many image fusion approaches in the remote sensing have been developed, including the intensity-hue-saturation (IHS) transform[1], principal component analysis (PCA)[2] and high pass filtering (HPF)[3], and so on. Recently, with the development of the wavelet transform, the discrete wavelet transform (DWT)[4] and the separable discrete wavelet frame (DWF)[5] transform have become powerful tools for multiscale image fusion. When the separable dyadic 2D wavelet is used to process images, it only applies the dyadic 1D wavelet separately to the rows and columns of the images. However, non-separable wavelet allows true processing of images. Images are treated as areas instead of rows and columns. The advantage with non-separable wavelet is having better frequency characteristic, directional property and more degree of freedom, resulting in better design[6].

In this paper, an approach based on discrete non-separable wavelet frame (NWF) is developed for combine high and low-resolution data. Experimental results indicate that the proposed method outperforms the approaches based on IHS transform, DWT and separable DWF transform.

The application of non-separable wavelet decomposition is similar to the 1D case. The low-pass component is repeatedly filtered and subsampled resulting in another low-pass and another detail signal. However, subsampling is not performed by retaining every second column and row, as it is in the separable case, but samples on a lattice are retained. There are many kinds of sampling lattices. In this paper, only the quincunx case is considered. One of the quincunx subsampling matrixes is given by

$$D_q = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{1}$$

The determinant of the matrix $D_q$ is two, so quincunx wavelet will correspond to a critically sampled two-channel filter bank. In each channel, the number of samples is reduced by the factor 2 and the horizontal and vertical directions are reduced by the scaling factor $\sqrt{2}$. This means that this new analysis will be twice as fine as the dyadic multiresolution analysis[7]. This is why we prefer to use the quincunx wavelet transform.

The scaling function of the quincunx wavelet is defined by

$$\phi_{m,n}(x,y) = 2^{-m}\phi\left(\zeta^{-2m}(x,y) - n\right), \quad n \in Z^2, \ \ n = (n_x, n_y), \tag{2}$$

where $\zeta(x,y) = (x+y, x-y)$ is a linear transform with $m \in \frac{1}{2}Z$.

The design of 2D non-separable filters is generally more difficult than the design of 1D filters. McClellan transformation has been shown to be a useful technique for designing 2D non-separable filters. The method consists of two parts: design of the 1D filters and the transformation. It allows to transform 1D prototype filter into 2D zero phase FIR filters, and the 2D filter parameterized by the McClellan transformation has the property of 1D prototype filter.

The Fourier transform of 1D zero phase symmetric filter can be expressed as a function of $\cos\omega$. The McClellan transform is to substitute the transformation $F(\omega_1, \omega_2)$ for $\cos\omega$ and yield a 2D zero phase filter. Often the following transform is chosen

$$F(\omega_1, \omega_2) = \frac{1}{2}(\cos\omega_1 + \cos\omega_2). \tag{3}$$

Compared with the non-separable wavelet transform, the overcomplete wavelet frame has two advantages: less constraint on filters and translation invariance. In each decomposition level, the NWF decomposition involves convolutions by insertion of an appropriate number of zeros into low-pass and high-pass filters and skipping the downsampling. So each frequency band has the same size.

The proposed image fusion approach using NWF consists of the following steps:

1) The low-resolution multispectral image is registered onto the high-resolution panchromatic image. And the

low-resolution multispectral image is resampled to yield the same pixel size as the high-resolution panchromatic image.

2) The R, G and B bands of the multispectral image are transformed into the IHS components and histogram matching between the panchromatic image and the intensity component of the multispectral image is performed.

3) The panchromatic image $f$ is decomposed into NWF representation: $PD_{i+1}(x)$ and $PS_{i+1}(x)$ with the same resolution.

$$\begin{aligned} PD_{i+1}(x) &= [g]_{\uparrow 2^i} * PS_i(x), \\ PS_{i+1}(x) &= [h]_{\uparrow 2^i} * PS_i(x), \\ &(i = 0, \cdots, N), \end{aligned} \tag{4}$$

where $PS_0 = f$; $h$ and $g$ are prototype filters and $[g]_{\uparrow 2^i}$ and $[h]_{\uparrow 2^i}$ are their dilated versions.

4) In order to improve the spatial information of the fused image, multiscale contrast enhancement method[8] is applied to enhance the panchromatic image. Every level band-pass image is mapped by a nonlinear function.

$$PE_i(x) = \begin{cases} G_i \cdot \left(1 - \frac{|x|}{M}\right)^p + x & \text{for } |x| \le M \\ x & \text{otherwise} \end{cases}, \tag{5}$$

where $M$ is the upper limit for the nonlinear enhancement, $G_i$ is a gain factor, $i$ is the number of level.

5) The approximated band of the panchromatic image is replaced by the intensity component of the multispectral image directly so that the spectral information of the multispectral image is preserved effectively. Low-frequency coefficients of the fused image are $\{I\}$ and high-frequency coefficients are $\{PE_i,\ i = 1, 2, \cdots, N\}$.

6) The final fused image is obtained by performing the inverse NWF transform and inverse IHS transform.

$$FZ = \tilde{h}_I * I + \sum_{i=1}^{I} \tilde{g}_i * PE_i(x). \tag{6}$$

Figure 1 illustrates the block diagram of the image fusion scheme using NWF. We assume that the number of decomposition levels is 2.

In this paper, we applied the proposed method to merge SPOT (10 m) panchromatic image and the corresponding Landsat TM (30 m) multispectral image of the London area, acquired in 1994. After registration between the SPOT image and TM image, the TM image is resampled by using the cubic convolution to the same size as the SPOT image. We select the (9,7)-taps biorthogonal filter as the 1D prototype filter of designing the non-separable filter. The parameters of multiscale enhancement are chosen as: $G_1 = 22.5$, $G_2 = 15$, $G_3 = 8$, $p = 3.0$, and $M = 3.5$.

For comparison with other fusion methods, the fusion scheme is also performed based on the IHS transform, DWT and DWF transform. The (9,7)-taps biorthogonal filters are also used for DWT and DWF transform.

In this paper, three evaluation criteria are used for quantitatively assessing the performance of the fusion results.

The average gradient of the image $f$ is

$$g = \frac{1}{(M-1)(N-1)} \times \sum_{x=1}^{M-1} \sum_{y=1}^{N-1} \sqrt{\frac{\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2}{2}}. \tag{7}$$

The average gradient reflects the clarity of an image. It can be used to measure the spatial resolution of the fused image. A larger average gradient means better spatial resolution.

The relative deviation between the fused image and the multispetral image $L$ is

$$DI = \frac{1}{M \cdot N} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{|F(i,j) - L(i,j)|}{L(i,j)}. \tag{8}$$

A smaller relative deviation implies better image quality.

The correlation coefficient between the fused image $F$ and the multispetral image $L$ is

$$\text{corr}(F/L) = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (F(i,j) - \overline{F})(L(i,j) - \overline{L})}{\sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (F(i,j) - \overline{F})^2 \sum_{i=1}^{m} \sum_{j=1}^{n} (L(i,j) - \overline{L})^2}}, \tag{9}$$

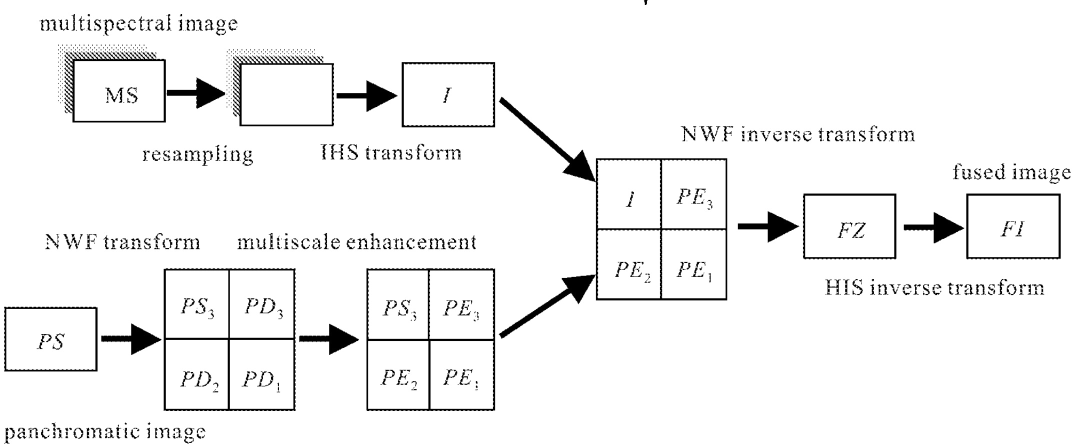


Fig. 1. Diagram of image fusion scheme using NWF.

where $\overline{F}$ and $\overline{L}$ are the mean values of the corresponding data set. A higher correlation value between the multispectral image and the fused image means that more spectral information in the multispectral image has been preserved.

The SPOT (10 m) panchromatic image and the corresponding Landsat TM (30 m) multispetral image are shown in Figs. 2(a) and (b), respectively. Figure 2(c) shows the fused image by using the IHS transform. The fusion result indicates the IHS transform produces evident spectral degradation though the spatial resolution is increased. Fusion results by using DWT, DWF and NWF transform are shown in Figs. 2(d), (e) and (f), respectively. The evaluation results of various fused images are compared in Table 1. Table 1 shows the average gradients of the fusion results by using various methods are larger than the average gradient of Landsat TM image. This means that various fusion methods can increase the spatial resolution of the fused images. However, the average gradient of NWF method is larger than of IHS, DWT and DWF methods, which indicates NWF method is capable of acquiring the better spatial quality. The second and third columns in Table 1 show the relative deviation and the correlation between the fused image and the Landsat TM image. Note that NWF yields the smallest relative deviation and the highest correlation coefficient. This means that NWF method preserves more spectral information than other fusion methods.

In the second experiment, the Landsat TM image in Fig. 2(a) is shifted artificially one pixel to the left. So

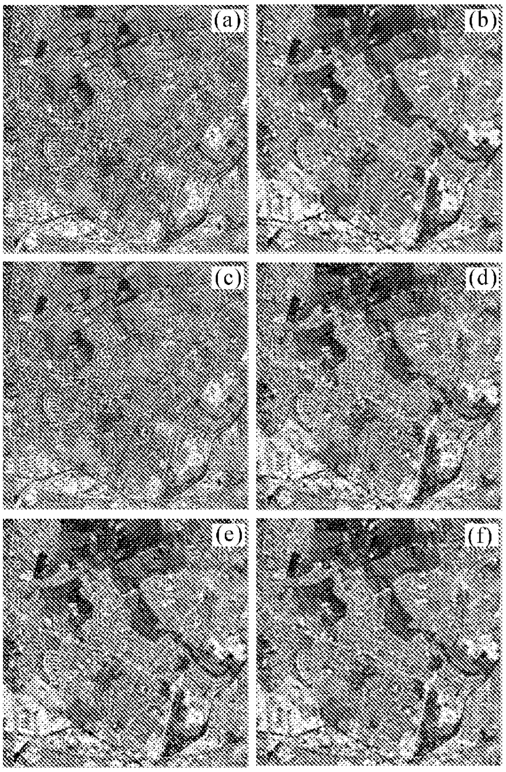


Fig. 2. The original images and the fused images. (a) The original SPOT (10 m) image; (b) the original Landsat TM (30 m) image; (c) the fused image by HIS transform; (d) the fused image by DWT; (e) the fused image by DWF transform; (f) the fused image by NWF transform.

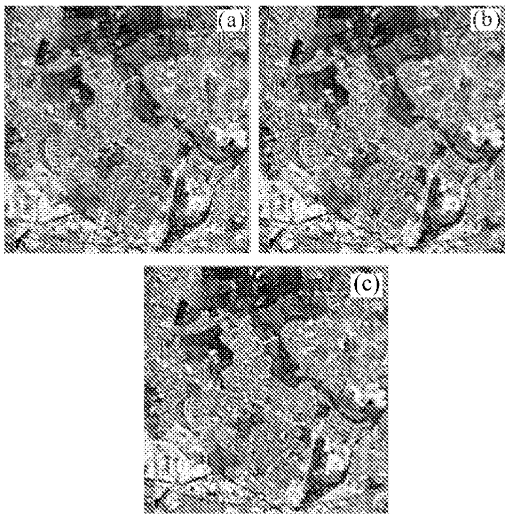


Fig. 3. (a) The fused image by DWT; (b) the fused image by DWF transform; (c) the fused image by NWF transform.

**Table 1. Evaluation Results of Fused Images**

| | Average Gradient | | | Relative Deviation | | | Correlation Coefficient | | |
|---|---|---|---|---|---|---|---|---|---|
| | R | G | B | R | G | B | R | G | B |
| IHS | 31.9957 | 31.8507 | 32.0872 | 0.5691 | 0.3951 | 0.5264 | 0.4995 | 0.3826 | 0.4762 |
| DWT | 31.4639 | 31.5187 | 31.6703 | 0.3537 | 0.2774 | 0.3409 | 0.7740 | 0.7052 | 0.7568 |
| DWF | 31.3861 | 31.4568 | 31.5866 | 0.3326 | 0.2653 | 0.3237 | 0.8009 | 0.7406 | 0.7857 |
| NWF | 33.1295 | 33.2004 | 33.3315 | 0.2983 | 0.2344 | 0.2905 | 0.8534 | 0.8074 | 0.8405 |
| TM | 10.1505 | 9.3723 | 9.9521 | | | | | | |

**Table 2. Evaluation Results of Fusion Images**

| | Average Gradient | | | Relative Deviation | | | Correlation Coefficient | | |
|---|---|---|---|---|---|---|---|---|---|
| | R | G | B | R | G | B | R | G | B |
| DWT | 31.3749 | 31.4726 | 31.6629 | 0.3713 | 0.2900 | 0.3544 | 0.7511 | 0.6757 | 0.7345 |
| DWF | 31.2985 | 31.4002 | 31.5765 | 0.3474 | 0.2751 | 0.3347 | 0.7819 | 0.7159 | 0.7673 |
| NWF | 33.0251 | 33.1121 | 33.2686 | 0.3039 | 0.2384 | 0.2948 | 0.8443 | 0.7951 | 0.8315 |

the original image is not perfectly registered. The corresponding fusion results by using DWT, DWF and NWF transform are shown in Figs. 3(a), (b) and (c), respectively. The evaluation results of various fused images are compared in Table 2.

In this paper, an image fusion method based on NWF transform is proposed for merging a high-resolution panchromatic image and a low-resolution multispectral image. The advantages of non-separable wavelet are having better frequency characteristic, directional property and more degree of freedom. Experiment results show that the proposed approach outperforms method based on the IHS transform, DWT and DWF transform. However, it possesses more memory space and computation time than others. Note that the fusion performance of various fusion methods deteriorates but the performance of the other fusion methods do more rapidly than NWF.

This research was supported by the National '863' Project of China (No. 2001AA135091), the Shanghai Key Project (No. 02DZ15001) and the Aviation Science Foundation of China (No. 02D57003). H. Wang's e-mail address is wangh@sjtu.edu.cn.

## References

1. J. W. Carper, T. M. Lillesand, and R. W. Kiefer, Photogrammetric Engineering and Remote Sensing **56,** 459 (1990).
2. P. S. Chavez, S. C. Slides, and J. A. Anderson, Photogrammetric Engineering and Remote Sensing **57,** 295 (1991).
3. V. K. Sheffigara, Photogrammetric Engineering and Remote Sensing **58,** 561 (1992).
4. J. Zhou, D. L. Civco, and J. A. Silander, International J. Remote Sensing **19,** 743 (1998).
5. S. T. Li, T. K. Ianes, and Y. N. Wang, International J. Information Fusion **3,** 17 (2002).
6. G. Karlsson and M. Vetterli, IEEE Trans. on Acoustics Speech and Signal Processing **38,** 925 (1990).
7. M. Barlaud, P. Sole, T. Gaidon, M. Antonini, and P. Mathieu, IEEE Trans. on Image Processing. **3,** 367 (1994).
8. M. Stahl, T. Aach, T. M. Buzug, S. Dippel, and U. Neitzel, Proc. SPIE **3661,** 1406 (1999).