ELSEVIER

Information Fusion 3 (2002) 17–23

INFORMATION FUSION

www.elsevier.com/locate/inffus

Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images

Shutao Li [a,b,*], James T. Kwok [a], Yaonan Wang [a]

[a] College of Electrical and Information Engineering, Hunan University, Changsha, China
[b] Department of Computer Science, Hong Kong University of Science and Technology, Hong Kong, China

Received 18 July 2000; received in revised form 11 June 2001; accepted 18 June 2001

Abstract

In this paper, we propose a pixel level image fusion algorithm for merging Landsat thematic mapper (TM) images and SPOT panchromatic images. The two source images are first decomposed using the discrete wavelet frame transform (DWFT), which is both aliasing free and translation invariant. Wavelet coefficients from TM's approximation subband and SPOT's detail subbands are then combined together, and the fused image is reconstructed by performing the inverse DWFT. Experimental results show that the proposed approach outperforms methods based on the intensity-hue-saturation transform, principal component analysis and discrete wavelet transform in preserving spectral and spatial information, especially in situations where the source images are not perfectly registered.

Keywords: Multisensor fusion; Wavelet; Wavelet frame; Remote sensing; Image processing

1. Introduction

Image fusion combines complementary information from multiple image sources for improved information or decision making [1–3]. Depending on the stage at which fusion takes place, it is often divided into three categories, namely, pixel level, feature level and decision level [4]. In pixel level fusion, the combination mechanism works directly on the pixels obtained at the sensors' outputs. Feature level fusion, on the other hand, works on image features extracted from the source images. Decision level fusion works at an even higher level, and merges the interpretations of different images obtained after image understanding. In this paper, we will focus on pixel level image fusion.

In remote sensing applications, fusion of a high spatial resolution image and a lower resolution multispectral image is often an efficient and economical means to produce a high spatial resolution multispectral image. As an example, Landsat thematic mapper (TM) data have six reflective 30 m resolution bands, while the *System Pour l'Observation de la Terre* (SPOT) data have 10 m resolution but is panchromatic. By fusing the Landsat TM and SPOT PAN images, we may be able to obtain both high spatial resolution and multispectral information in one single image.

Commonly used data fusion methods in the remote sensing community include the intensity-hue-saturation (IHS) transform [5], principal component analysis (PCA) [6], high pass filtering [7], and regression variable substitution [8]. Recently, the discrete wavelet transform (DWT) has also been used [9–11]. However, because of an underlying down-sampling process in DWT, its multiresolution decompositions and consequently the fusion results are shift variant. This is particularly undesirable when the source images cannot be perfectly registered.

In this paper, we propose an image fusion scheme for Landsat TM and SPOT PAN images based on the discrete wavelet frame transform (DWFT) [12,13]. DWFT avoids down-sampling by using an overcomplete wavelet decomposition. Consequently, it is both aliasing free

* Corresponding author. Tel.: +852-2358-7027; fax: +852-2358-1477.
E-mail addresses: shutaoli@cs.ust.hk, shutao_li@yahoo.com.cn (S. Li), jamesk@cs.ust.hk (J.T. Kwok), yaonan@mail.hunu.edu.cn (Y. Wang).

1566-2535/02/$ - see front matter
PII: S1566-2535(01)00037-9

and translation invariant. After the decomposition, wavelet coefficients from TM's approximation subband and SPOT's detail subbands are combined together, and the fused image is then obtained by performing the inverse DWFT. Experimental comparison with IHS, PCA and DWT shows that the proposed method outperforms these existing approaches, especially in situations where the source images are not perfectly registered.

The rest of this paper is organized as follows. Introduction to a number of common data fusion methods, including the IHS, PCA and DWT, is given in Section 2, followed by a brief review of the DWFT in Section 3. Section 4 describes our proposed fusion scheme for Landsat TM and SPOT PAN images. Experimental results are presented in Section 5, and Section 6 gives some concluding remarks.

2. Common data fusion methods in remote sensing

2.1. Intensity-hue-saturation (IHS) transform

IHS is one of the commonest fusion methods for remote sensing data. The three multispectral bands, R, G and B, of a low resolution image are first transformed to the IHS color space as [5]

$$\begin{pmatrix} I \\ V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}, \tag{1}$$

$$H = \tan^{-1}\left(\frac{V_2}{V_1}\right), \tag{2}$$

$$S = \sqrt{V_1^2 + V_2^2}. \tag{3}$$

Here I is the intensity, H is the hue, S is the saturation, and V_1, V_2 are the intermediate variables. Fusion proceeds by replacing I with the panchromatic high-resolution image information. The fused image is then obtained by performing an inverse transformation from IHS back to the original RGB space as

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ 1 & \frac{1}{\sqrt{6}} & -\frac{1}{2} \\ 1 & -\frac{2}{\sqrt{6}} & 0 \end{pmatrix} \begin{pmatrix} I \\ V_1 \\ V_2 \end{pmatrix}. \tag{4}$$

2.2. Principal component analysis (PCA)

PCA is a general statistical technique that transforms multivariate data with correlated variables into one with uncorrelated variables. These new variables are obtained as linear combinations of the original variables. PCA has been widely used in image encoding, image data compression, image enhancement and image fusion. When used in image fusion, it is performed on the image with all its spectral bands [6]. The first principal component is then replaced with the panchromatic high-resolution image information, and the fused image is obtained by inverting back to the original color space.

2.3. Discrete wavelet transform (DWT)

DWT decomposes a signal into a multiresolution representation with both low frequency coarse information and high frequency detail information. When used for image fusion, the source images are first geometrically registered, and then decomposed by DWT to the same resolution. Corresponding wavelet coefficients are combined, and the fused image is obtained by performing the inverse wavelet transform. Since the distributions of coefficients in the detail subbands have mean zero, the fusion result does not change the radiometry of the original multispectral image.

3. Discrete wavelet frame transform (DWFT)

The biorthogonal DWFT uses two different wavelet bases: $\Psi(x)$ for decomposition and $\tilde{\Psi}(x)$ for reconstruction. As shown in Fig. 1, the 2-D DWFT is implemented using a bank of 1-D low pass ($h(x)$) and high pass ($g(x)$) analysis filters. The reconstruction process is similarly computed via 1-D synthesis filters, $\tilde{h}(x)$ and $\tilde{g}(x)$. To obtain a perfect reconstruction, these filters must satisfy

$$H(\omega)\tilde{H}(\omega) + G(\omega)\tilde{G}(\omega) = 1, \tag{5}$$

$$g(x) = (-1)^{x-1}\tilde{h}(1-x), \tag{6}$$

$$\tilde{g}(x) = (-1)^{x-1}h(-1-x). \tag{7}$$

Furthermore, the analysis filters can be expressed as

$$H(\omega) = \frac{\tilde{\Phi}(2\omega)}{\Phi(\omega)}, \tag{8}$$

$$G(\omega) = \frac{\tilde{\Psi}(2\omega)}{\Psi(\omega)}, \tag{9}$$

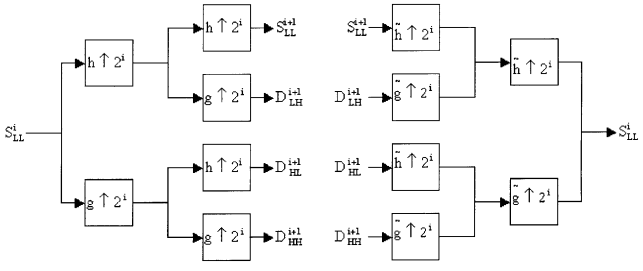

Fig. 1. One stage of 2-D DWFT and inverse DWFT.

where $\tilde{\Phi}(\omega)$ and $\tilde{\Psi}(\omega)$ are the Fourier transforms of the scaling function $\tilde{\Phi}(x)$ and synthesis wavelet $\tilde{\Psi}(x)$, respectively.

In the decomposition phase of 2-D DWFT, each row of the image is separately filtered by H and G. The resulting row-transformed image is then similarly filtered in the column direction, finally yielding four subbands at the first decomposition level ($j = 1$). The three detail subbands, D^j_{LH}, D^j_{HL}, D^j_{HH}, contain the vertical, horizontal, and diagonal high frequency information, respectively, while the approximation subband S^j_{LL} is a low pass filtered version of the original image. This approximation subband is subsequently passed to the next level for further subband decomposition. Thus, a DWFT with N decomposition levels will have a total of $3N + 1$ frequency subbands, all of them are of the same size.

4. Fusion with DWFT

The proposed image fusion algorithm (Fig. 2) consists of the following steps:

1. The Landsat TM image is geometrically registered onto the SPOT PAN image. The Landsat TM image is super-sampled to 10 m resolution so that both the SPOT PAN and TM images occupy the same geographic space and possess the same pixel size.
2. Both images are decomposed into DWFT representation with the same resolution. For simplicity of illustration, we here assume that the number of decomposition levels is 2. Denote the decomposed subbands of the SPOT PAN image and (one band of) the Landsat TM image by $\{SD^1_{\mathrm{LH}}, SD^1_{\mathrm{HL}}, SD^1_{\mathrm{HH}}, SD^2_{\mathrm{LH}}, SD^2_{\mathrm{HL}}, SD^2_{\mathrm{HH}}, SS^2_{\mathrm{LL}}\}$ and $\{MD^1_{\mathrm{LH}}, MD^1_{\mathrm{HL}}, MD^1_{\mathrm{HH}}, MD^2_{\mathrm{LH}}, MD^2_{\mathrm{HL}}, MD^2_{\mathrm{HH}}, MS^2_{\mathrm{LL}}\}$, respectively.
3. Merge the approximation subband of the Landsat TM image and the detail subbands of the SPOT PAN image into a fused DWFT representation. Denote the decomposed subbands in the fused DWFT representation by $\{FD^1_{\mathrm{LH}}, FD^1_{\mathrm{HL}}, FD^1_{\mathrm{HH}}, FD^2_{\mathrm{LH}}, FD^2_{\mathrm{HL}}, FD^2_{\mathrm{HH}}, FS^2_{\mathrm{LL}}\}$. Then,

$$FS^2_{\mathrm{LL}} = MS^2_{\mathrm{LL}}, \tag{10}$$

$$FD^1_{\mathrm{LH}} = SD^1_{\mathrm{LH}}, \tag{11}$$

$$FD^1_{\mathrm{HL}} = SD^1_{\mathrm{HL}}, \tag{12}$$

$$FD^1_{\mathrm{HH}} = SD^1_{\mathrm{HH}}, \tag{13}$$

$$FD^2_{\mathrm{LH}} = SD^2_{\mathrm{LH}}, \tag{14}$$

$$FD^2_{\mathrm{HL}} = SD^2_{\mathrm{HL}}, \tag{15}$$

$$FD^2_{\mathrm{HH}} = SD^2_{\mathrm{HH}}. \tag{16}$$

4. Perform the inverse DWFT on the fused DWFT representation to obtain the final fused image.

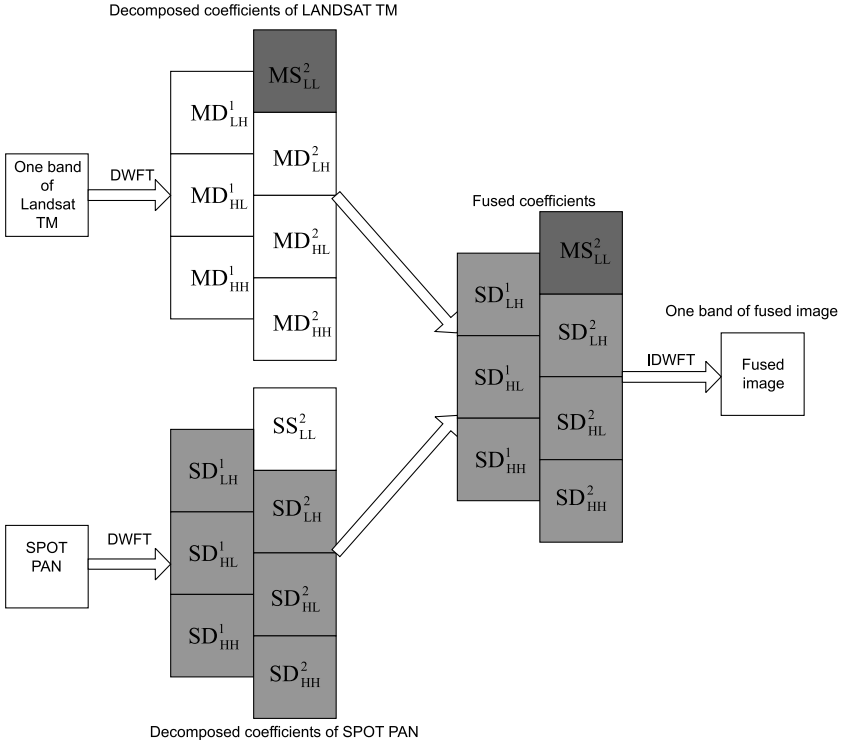

Fig. 2. Schematic diagram for image fusion using DWFT.

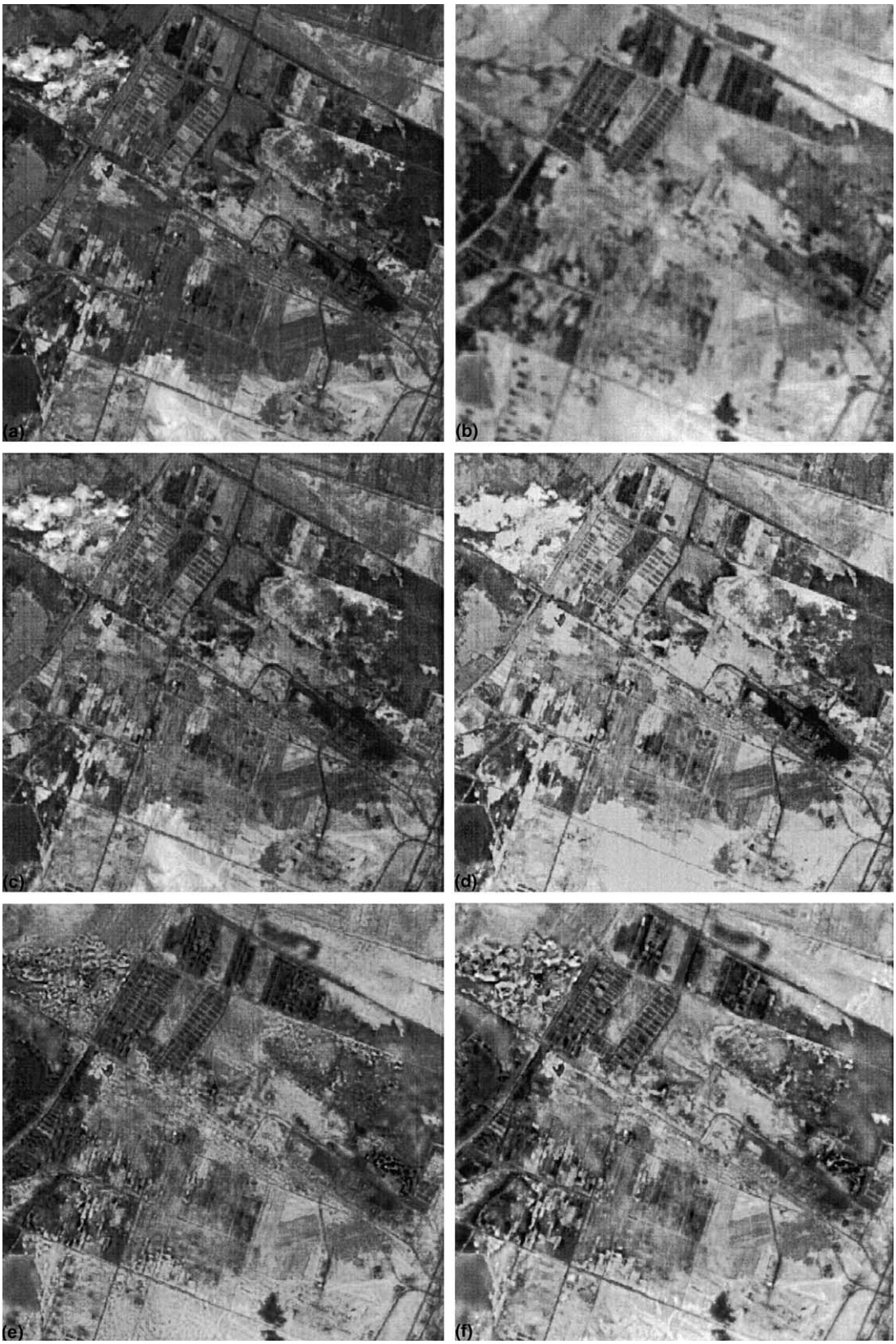

Fig. 3. The original images and fusion results: (a) the original SPOT image; (b) the original Landsat TM image; (c) fusion result by IHS; (d) fusion result by PCA; (e) fusion result by DWT; (f) Ffsion result by DWFT.

5. Experiments

5.1. Setup

Fig. 3(a) shows a SPOT PAN image of the Ningxia area in western China, acquired on February 17, 1993. Fig. 3(b) shows the corresponding Landsat TM image, acquired on May 28, 1995, after resampling and registration. The six Landsat TM bands $(B_1, B_2, B_3, B_4, B_5, B_7)$ are converted to the (R, G, B) color space by

$$R = \frac{B_5 + B_7}{2}, \quad G = \frac{B_3 + B_4}{2} \quad \text{and} \quad B = \frac{B_1 + B_2}{2}.$$

We use the biorthogonal B-spline function [14], together with a decomposition level of 3, as the basis for DWFT. The corresponding filter coefficients for $h(x)$ and $g(x)$ are listed in Table 1. For comparison purposes, fusion is also performed on using the IHS transform, PCA and DWT. We use the Daubechies' db8, also with a decomposition level of 3, as the wavelet basis for DWT.

Fusion should preserve both the spectral characteristics of the high spectral resolution image and the spatial characteristics of the high spatial resolution image. The spectral quality of an $M \times N$ fused image can be measured by the discrepancy D_k at each band

$$D_k = \frac{1}{MN} \sum_i \sum_j \left| F_{k,i,j} - L_{k,i,j} \right|, \tag{17}$$

where $F_{k,i,j}$ and $L_{k,i,j}$ are the pixel values at position (i, j) in the kth band of the fused and original Landsat TM images, respectively. For the spatial quality, we notice that spatial information unique to the SPOT PAN image is mostly concentrated in the high frequency domain. Hence, the correlation coefficient between the high pass component of the fused image and that of the SPOT PAN image can be used as a measure, with a higher correlation implying that more spatial information in the SPOT PAN image has been preserved. The following Laplacian filter

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

is used for extracting the high pass component.

5.2. Results

Fusion results by using the IHS transform and PCA are shown in Figs. 3(c) and (d), respectively. The increase in resolution with respect to the original Landsat TM image is evident. However, both methods suffer from the problem that the radiometry on the spectral channels is modified after fusion. This is because the high-resolution panchromatic image usually has spectral characteristics different from both the intensity image I and the first principal component. Fusion results on using DWT and the proposed DWFT methods are shown in Figs. 3(e) and (f), respectively. The spectral and spatial qualities of various fusion results are compared in Table 2. DWFT yields the smallest spectral difference and highest spatial correlation. The performance of DWT is also quite competitive in this case.

In the second experiment, we artificially shift the Landsat TM image in Fig. 3(b) by one pixel to the right to obtain Fig. 4(a). The source images are thus not perfectly registered. Results on fusing Fig. 3(a) and Fig. 4(a) are shown in Fig. 4 and Table 3. As can be seen, the performance of DWT deteriorates rapidly and the fusion result by DWFT is much superior.

Table 1
Coefficients for the analysis (h) and synthesis ($\tilde{h}$) filters in the biorthogonal B-spline wavelet basis

N	0	±1	±2	±3	±4
h	0.602949	0.266864	−0.078223	−0.016864	0.026749
$\tilde{h}$	0.557543	0.295636	−0.028772	−0.045636	0

Table 2
Spectral and spatial qualities of the fusion results

	Spectral discrepancy			Spatial correlation		
	R	G	B	R	G	B
IHS	32.5059	26.0224	26.7658	0.9183	0.9403	0.9465
PCA	18.3211	18.3149	18.2481	0.9299	0.9424	0.9563
DWT	15.3483	12.7571	12.0803	0.9812	0.9931	0.9933
DWFT	13.6436	11.9334	11.3622	0.9840	0.9945	0.9957

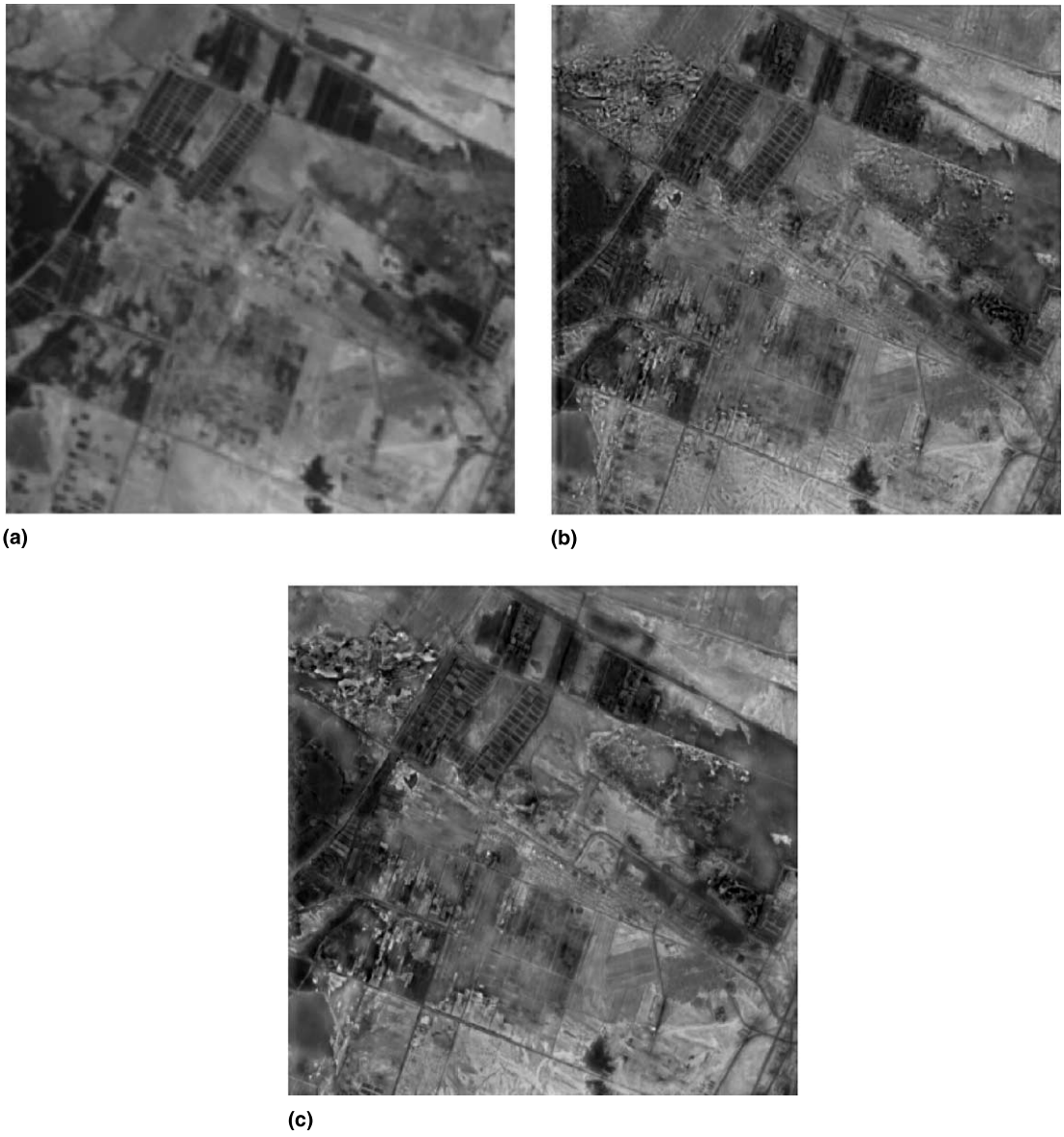

Fig. 4. The shifted Landsat TM image and fusion results: (a) the shifted Landsat TM image; (b) fusion result by DWT; (c) fusion result by DWFT.

Table 3
Spectral and spatial qualities of the fusion results based on the shifted Landsat TM image

	Spectral discrepancy			Spatial correlation		
	R	*G*	*B*	*R*	*G*	*B*
DWT	16.1738	13.4293	12.9680	0.9793	0.9909	0.9912
DWFT	14.5679	12.6654	12.3260	0.9815	0.9916	0.9928

6. Conclusion

The integration of complementary spectral and spatial remote multisensor data can facilitate visual and automatic interpretation. In this paper, we propose fusing SPOT PAN and Landsat TM images by using the DWFT. Because DWFT has no dyadic decimation at each decomposition level, its representation is both aliasing free and translation invariant. Experimental results show that DWFT outperforms methods based on IHS, PCA and DWT, especially when the source images are not perfectly registered. A disadvantage of DWFT, however, is that it is more demanding in terms of both memory and time.

Acknowledgements

The authors would like to thank the anonymous reviewers for their constructive comments on an earlier version of this paper.

References

[1] D.L. Hall, J. Llinas, An introduction to multisensor data fusion, Proceedings of the IEEE 85 (1) (1997) 6–23.
[2] P.K. Varshney, Multisensor data fusion, Electronics and Communication Engineering Journal 9 (12) (1997) 245–253.
[3] J.K. Aggarwal, Multisensor Fusion for Computer Vision, Springer, Berlin, 1993.
[4] C. Pohl, Multisensor image fusion in remote sensing: concepts, methods and applications, International Journal of Remote Sensing 19 (5) (1998) 823–854.
[5] W.J. Carper, T.W. Lilesand, R.W. Kieffer, The use of Intensity-Hue-Saturation transformation for merging SPOT panchromatic and multispectral image data, Photogrammetric Engineering and Remote Sensing 56 (4) (1990) 459–467.
[6] P.S. Chavez, S.C. Sildes, J.A. Anderson, Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic, Photogrammetric Engineering and Remote Sensing 57 (3) (1991) 295–303.
[7] V.K. Sheffigara, A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set, Photogrammetric Engineering and Remote Sensing 58 (5) (1992) 561–567.
[8] A. Singh, Digital change detection techniques using remotely sensed data, International Journal of Remote Sensing 10 (6) (1989) 989–1003.
[9] H. Li, B.S. Manjunath, S.K. Mitra, Multisensor image fusion using the wavelet transform, Graphical Models and Image Processing 57 (3) (1995) 235–245.
[10] D.A. Yockey, Image merging and data fusion by means of the discrete two-dimensional wavelet transform, Journal of Optical Society of America A 12 (9) (1995) 1834–1841.
[11] J. Zhou, D.L. Civco, J.A. Silander, A wavelet transform method to merge Landsat TM and SPOT panchromatic data, International Journal of Remote Sensing 19 (4) (1998) 743–757.
[12] M. Unser, Texture classification and segmentation using wavelet frames, IEEE Transactions on Image Processing 4 (11) (1995) 1549–1560.
[13] A. Laine, J. Fan, Frame representations for texture segmentation, IEEE Transactions on Image Processing 5 (5) (1996) 771–780.
[14] A.I. Cohen, I. Daubechies, J.C. Feauveau, Biorthogonal bases of compactly supported wavelets, Communications on Pure and Applied Mathematics 45 (5) (1992) 485–560.