FUSING SAR AND OPTICAL IMAGES BASED ON COMPLEX WAVELET TRANSFORM

Shuai Xing*, Qing Xu

Dept. of Remote Sensing Information Engineering, Institute of Surveying and Mapping, 450052 Longhai Middle Road, Zhengzhou City, China-xing972403@163.com

Commission VII, WG VI

KEY WORDS: Algorithms, Multiresolution analysis, Synthetic aperture radar, Image fusion, Multispectral, Remote sensing

ABSTRACT:

Image fusion deals with the integration of remote sensing images from various sensors, such as SAR and optical images, aiming at achieving improved image information to better support improved image classification, monitoring and etc. The main goal of this paper is to introduce an algorithm to fuse SAR and multi-spectral optical images based on complex wavelet. First, the theoretical basis of complex wavelet is described together with its key properties (e.g. approximate shift invariance, good directional selectivity, perfect reconstruction (PR),limited redundancy and efficient order-N computation). Secondly, the fusing algorithm based on complex wavelet is proposed, which includes a method to de-noise SAR image and a new fusion rule based on modulus of complex wavelet coefficients. Finally, experiment results show that the fusion method based on dual-tree complex wavelet transform (DT-CWT) is remarkably better than that based on discrete wavelet transform (DWT).

1. INRTRODUCTION

SAR and optical remote sensing image fusion is aiming at achieving improved image quality to better support improved image classification, monitoring and etc. Fused image will enhance reliability and speed of feature extraction, increase the usage of the data sets, and extend remote sensing images' application area. There have been a lot of research efforts on image fusion, and many fusion methods have been proposed. One of them, the fusion algorithm based on DWT, has been canonized.

The advantages of wavelet are that it can analyze signal in time domain and frequency domain respectively and the multi-resolution analysis is similar with Human Vision System. DWT in maximally decimated form established by Mallat (S G Mallat, 1989) is widely used in image processing now, such as image matching, image segmentation, image classification, image fusion and so on. The best advantage of fusion based on DWT is to conserve more spectral characteristics of the multi-spectral image. So the fusion algorithm based on DWT is widely used. But DWT has two main disadvantages (N. Kingsbury, 1998a):

- Lack of shift invariance. This means that small shifts in the input signal can cause major variations in the distribution of energy between wavelet coefficients at different scales.
- Poor directional selectivity for diagonal features, because the wavelet features are separable and real.

Nick Kingsbury has introduced the dual-Tree complex wavelet transform (DT-CWT), which has the following properties (N. Kingsbury, 1998a):

- Approximate shift invariance;
- Good directional selectivity in 2-dimensions (2-D) with Gabor-like filters also true for higher dimensionality: m-D);
- Perfect reconstruction (PR) using short linear-phase filters;
- Limited redundancy: independent of the number of scales: 2:1 for 1-D (2^m :1 for m-D);
- Efficient order-N computation - only twice the simple DWT for 1-D (2^m times for m-D).

DT-CWT has shown good performance in image restoration and denoising (A. Jalobeanu , 2000; Nick Kingsbury, 1998b; Peter de Rivaz, 2001), motion estimation (Julian Magarey, 1998), image classification (Serkan Hatipoglu, 1999), texture analysis (Javier Portilla, 1999; N. Kingsbury, 1998; Serkan Hatipoglu, 1999), image enhancement (Nick Kingsbury , 1998b), image matching (JIANG Han-ping , 2000).

In this paper, we proposed an SAR and optical image fusion algorithm based on DT-CWT, and use a Radarsat-1 SAR image and a SPOT5 multi-spectral image to test the performance of our algorithm.

2. THE DUAL-TREE COMPLEX WAVELET TRANSFORM

It is well-known that the real biorthogonal wavelet transform can provide PR and no redundancy, but it is lack of shift variant. Then Kingsbury (Julian Magarey, 1998;N. Kingsbury, 1998a ; Nick Kingsbury, 1998b; Serkan Hatipoglu, 1999) has developed a dual-tree algorithm with a real biorthogonal transform, and an approximate shift invariance can be obtained by doubling the sampling rate at each scale, which is achieved by computing two parallel subsampled wavelet trees respectively.

For one dimension signal, we can compute two parallel wavelet trees. There is one sample offset delay between two trees at level 1, which is achieved by doubling all the sample rates. The shift invariance is perfect at level 1, since the two trees are fully decimated. To get uniform intervals between two trees beyond level 1, there have to be half a sample delay. The term will be satisfied using odd-length and even-length filters alternatively from level to level in each tree. Because we use the decimated form of a real discrete wavelet transform beyond level 1, the shift invariance is approximate.

The transform algorithm is described as following.

➢ At level 1, there is one sample offset between the trees.

$$(a_A^1)_n = (a^0 * h^O)_{2n} \qquad (d_A^1)_n = (a^0 * g^O)_{2n}$$
$$(a_B^1)_n = (a^0 * h^O)_{2n+1} \qquad (d_B^1)_n = (a^0 * g^O)_{2n+1} \tag{1}$$

➢ Beyond level 1, there must be half a sample difference between the trees.

$$(a_A^{j+1})_n = (a_A^j * h^e)_{2n} \qquad (d_A^{j+1})_n = (a_A^j * g^e)_{2n}$$
$$(a_B^{j+1})_n = (a_B^j * h^o)_{2n+1} \qquad (d_B^{j+1})_n = (a_B^j * g^o)_{2n+1} \tag{2}$$

The details d_A and d_B can be interpreted as the real and imaginary parts of a complex process $z = d_A + id_B$. The essential property of this transform is that the magnitude of the step response is approximately invariant with the input shift, while only the phase varies rapidly. (A. Jalobeanu , 2000)

Reconstruction is performed independently in each tree, and the results are averaged to obtain a^0 at level 1, for symmetry between the two trees.

- Level j $(j > 0)$:

$$(a_A^j)_n = (\bar{a}_A^{j+1} * \tilde{h}^e)_n + (\bar{d}_A^{j+1} * \tilde{g}^e)_n$$
$$(a_B^j)_n = (\bar{a}_B^{j+1} * \tilde{h}^o)_n + (\bar{d}_B^{j+1} * \tilde{g}^o)_n \tag{3}$$

- At $j = 0$:

$$a_n^0 = \frac{1}{2}((\bar{a}_A^1 * \tilde{h}^o)_n + (\bar{d}_A^1 * \tilde{g}^o)_n + (\bar{a}_B^1 * \tilde{h}^o)_n + (\bar{d}_B^1 * \tilde{g}^o)_n) \tag{4}$$

where $\bar{x}_n = \begin{cases} x_p & if \quad n = 2p \\ 0 & if \quad n = 2p+1 \end{cases}$, $\tilde{x}_n = \begin{cases} x_p & if \quad n = 2p+1 \\ 0 & if \quad n = 2p \end{cases}$

For 2-D signal, we can filter separately along columns and then rows by the way like 1-D. Kingsbury figured out in (Nick Kingsbury，1998a) that, to represent fully a real 2-D signal, we must filter with complex conjugates of the column and row filters. So it gives 4:1 redundancy in the transform. Furthermore, it remains computationally efficient, since actually it is close to a classical real 2-D wavelet transform at each scale in one tree, and the discrete transform can be implemented by a ladder filter structure.

The quad-tree transform is designed to be, as much as possible, translation invariant. It means that if we decide to keep only the details or the approximation of a given scale, removing all other scales, shifting the input image only produces a shift of the reconstructed filtered image, without aliasing. (A. Jalobeanu , 2000)

The most important property of CWT is that it can separate more directions than the real wavelet transform. The 2-D CWT can provide six subimages in two adjacent spectral quadrants at each level, which are oriented at angles of $\pm 15^\circ$, $\pm 45^\circ$, $\pm 75^\circ$. The strong orientation occurs because the complex filters are asymmetry responses. They can separate positive frequencies from negative ones vertically and horizontally. Figure 1 shows the transform of an isotropic synthetic image at level 3, which also contains details at different scales.

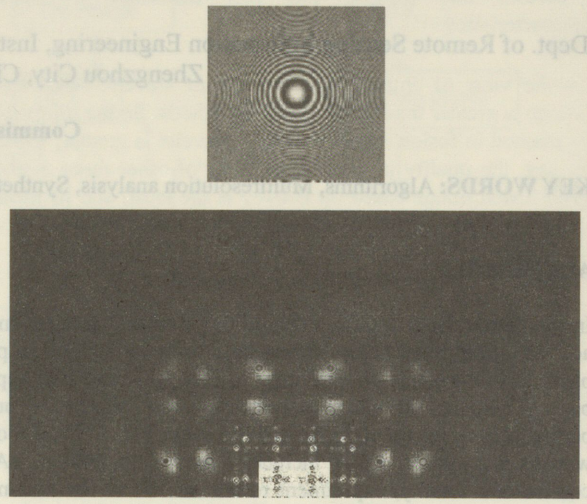

Figure 1. Isotropic test image containing various scale information (left), magnitude of its complex wavelet transform at level 3 showing both directional and scaling properties (right)

The DT-CWT is a good solution to image fusion because of its advantages. First, it is approximate shift invariant. If the input signal shifts a few samples, the fused image will be reconstructed without aliasing, which is useful to the not strictly registered images. Secondly, it can separate positive and negative frequencies and provide 6 subimages with different directions at each scale. So the details of DT-CWT can conserve more detail information than DWT. In addition, PR, limited redundancy and high computation efficiency make it suitable for image fusion execution.

3. SAR AND OPTICAL IMAGE FUSION BASED ON DT-CWT

3.1 Speckle Denoising

The SAR image is produced by coherently receiving echo. Echo overlapping inevitably produced speckle noise. Speckle is a serious obstacle of SAR image object recognition and makes some ground features disappear. (Xiao Guochao, 2001) So speckle has to be removed before image fusion.

A few algorithms, such as Lee, Frost, Kuan, Gamma MAP, are successfully used to denoise speckle with an assumption that speckle is multiplicative noise. Here the Lee-Sigma (Lee S. J., 1980) and Gamma MAP algorithms (Lopes A. et al. , 1993; Baraldi A. et al. , 1995) are chosen, because they can decrease the lost of edge features while removing speckle noise.

The Lee-Sigma algorithm is described as following

$$R = I + K \times (CP - U \times I) \tag{5}$$

$$K = 1 - \frac{(Sigma / U^2)}{(QVAR / I^2)} \tag{6}$$

Where

R is the grey level of the filtered interest pixel,
$QVAR$ is the variance in filter window,
I is the mean grey level in the filter window,
U is the mean multiplicative noise and usually is 1,
CP is the central pixel in filter window,
$Sigma$ is the multiplicative noise variance, it is estimated based on a Rayleigh distribution and consistent with those derived from actual data.

The Gamma MAP filter is based on a multiplicative noise model with non-stationary mean and variance parameters. Recent work has shown natural vegetated areas have been shown to be more properly modeled as having a Gamma distributed cross section. This algorithm incorporates this assumption. The exact formula used is:

$$R = \begin{cases} I & C_i \le C_u \\ (B \times I + \sqrt{D})/(2a) & C_u < C_i < C_{\max} \\ CP & C_i \ge C_{\max} \end{cases} \quad (7)$$

Where
$B = \alpha - NLOOK - 1$,
$D = I^2 \cdot B^2 + 4 \cdot \alpha \cdot NLOOK \cdot I \cdot CP$,
$\alpha = (1 + C_u^2)/(C_i^2 - C_u^2)$,
$C_u = 1/\sqrt{NLOOK}$,
$C_i = \sqrt{VAR}/I$,
$C_{\max} = \sqrt{2} * C_u$,
$NLOOK$ is number of looks,
VAR is variance in filter window.

By experiments we find using both Gamma MAP and Lee-Sigma filters to achieve better result than using Gamma MAP or Lee-Sigma filter twice. So here the SAR image is first filtered by Gamma MAP and then filtered by Lee-Sigma. The proportions of original SAR image and denoised SAR image are shown in figure 2. The speckle noise of denoised image has been obviously removed and edge features have been conserved.

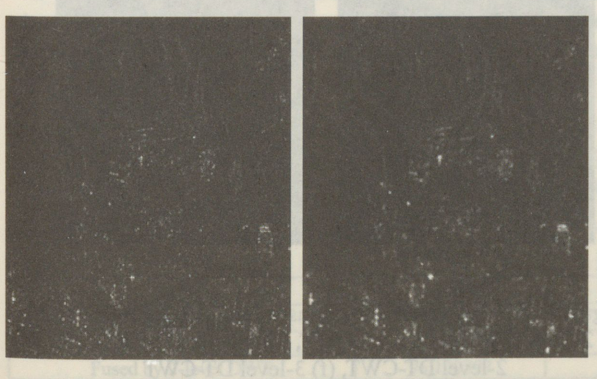

Figure 2. Proportions of original SAR image (left) and denoised SAR image (right)

3.2 The fusion algorithm

We design an algorithm based on DT-CWT for fusing a multi-spectral optical image and a SAR image. First the registered multi-spectral image and SAR image are decomposed by DT-CWT respectively, then the approximate and detail parts of two images are fused according to some rules at each level, finally the fused image is reconstructed. This procedure is illustrated by figure 3. The fusion procedure can be described in detail as following:

(1)Each band of the multi-spectral optical image and the SAR image are geometrically registered to each other. After geometrical rectification, their sizes are same.

(2)The gray level of SAR image is stretched tally with each band of multi-spectral images respectively using histogram equalization.

(3)Decompose the histogram-specified SAR and registered multi-spectral optical images with DT-CWT to form their multi-resolution and multi-directional descriptions. At the same time, the moduli of their complex wavelet transform are achieved.

(4)Since the aim of image fusion is to improve image information quality, we should analyze characteristics of SAR and optical images. Some objects, like lakes, roads or buildings, are distinct in SAR image but more details are hard to recognize. On the contrary, there are enough details and spectral information in optical image. So we design different fusion rules for low and high frequency parts fusion to integrate the advantages of two images.

Image fusion begins with the coarsest level. The gray value of a fused low frequency part pixel is determined by maximum gray value rule. The bigger absolute gray value at cooresponding pixel between SAR and optical images is selected. This rule makes more approximate parts and spectral information in optical image conserved.

The important information in SAR image is mostly in the high frequency parts. But some important details in optical image are also in the high frequency parts. So we decide to determine the fused pixel by comparing energy values of corresponding pixels in two images. The pixel with bigger energy value is the fused pixel. The energy value of a pixel is calculated in its centered neighbor window. Considering that DT-CWT of the images can be interpreted as a complex including real part and imaginary part, and the modulus can show clear directionality, the energy values can be computed according to the moduli of the high frequency parts. The procedure is illustrated in fig. 3.

The wavelet coefficients at point (i, j) of real and imaginary parts in the SAR image are denoted as $W_R^S(i, j)$ and $W_I^S(i, j)$ respectively. The wavelet coefficients at point (i, j) of real and imaginary parts in the optical image are denoted as $W_R^O(i, j)$ and $W_I^O(i, j)$ respectively. The magnitudes at point (i, j) in the SAR image and the optical image are achieved respectively by

$$M^S(i, j) = \sqrt{\left(W_R^S(i, j)\right)^2 + \left(W_I^S(i, j)\right)^2}$$
$$M^O(i, j) = \sqrt{\left(W_R^O(i, j)\right)^2 + \left(W_I^O(i, j)\right)^2} \quad (8)$$

The energy values at point (i, j) in the SAR image and the optical image are achieved respectively by

$$EN^S(i,j) = \sum_{k,l \in D(i,j)} [M^S(k,l) - E(M^S)]^2$$
$$EN^O(i,j) = \sum_{k,l \in D(i,j)} [M^O(k,l) - E(M^O)]^2 \quad (9)$$

Where $D(i,j)$ is the neighbor window of point (i,j). The window size of $D(i,j)$ usually is 3*3 or 5*5.

The fused wavelet coefficient $CW(i,j)$ at point (i,j) is obtained as following

$$CW(i,j) = \begin{cases} W^S(i,j) & EN^S(i,j) \geq EN^O(i,j) \\ W^O(i,j) & EN^S(i,j) < EN^O(i,j) \end{cases} \quad (10)$$

And then, the inverse DT-CWT is carried out for composing the new merged images at this level.

(5)The composing procedure in (4) are carried out recursively at their top levels until the first level is processed. This results in three new images.

(6) The three new produced images are compounded into one fused image. The fused image contains both the spectral information of multi-spectral optical images and the object structure information of SAR image.

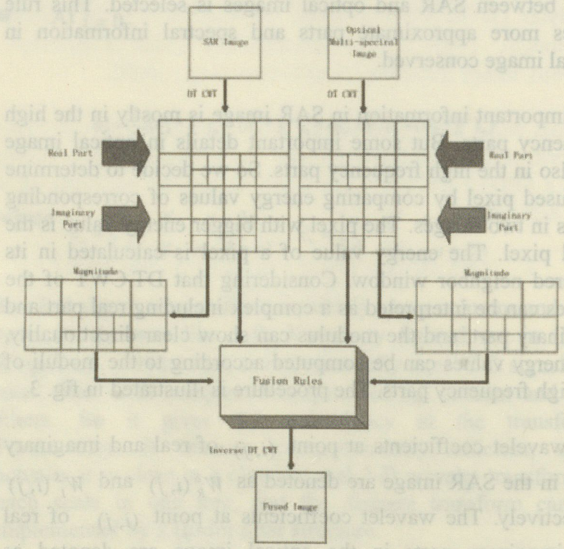

Figure 3. Procedure of image fusion based on DT-CWT

4. EXPERIMENTS

We chose two images in experiments. One is a Radarsat-1 SAR image (acquired in 6th November 2002, along-track slant range resolution is 8.82 meters, across-track slant range resolution is 5.56 meters) and a SPOT5 multi-spectral image composed of XS1, XS2 and XS3 bands (acquired in 1st October 2002, ground resolution is 10 meters). They are shown in figure 4. They have been registered strictly at the same scale. We fuse the images with DWT and DT-CWT at 1, 2, 3 level. Proportions of the fused images are listed in figure 5.

We find much more details of urban area have been lost for speckle noise in SAR image, but the lake at center and some buildings at right are distinct, because water reverberates less electromagnetic wave but the building produces strong echo. The buildings with strong echo are white points in the SAR image, and they clearly show us positions and shapes of the buildings and roads in the city. So these object features are what we need to fuse with multi-spectral optical image.

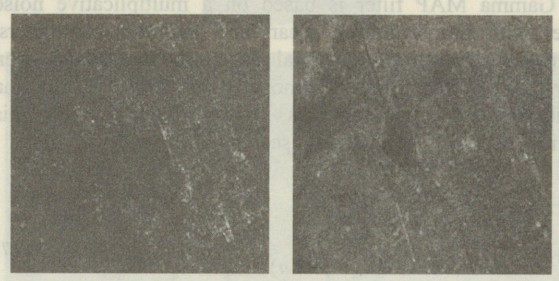

Figure 4. SAR image (left) and SPOT5 multi-spectral image composed of XS1, XS2 and XS3 bands (right)

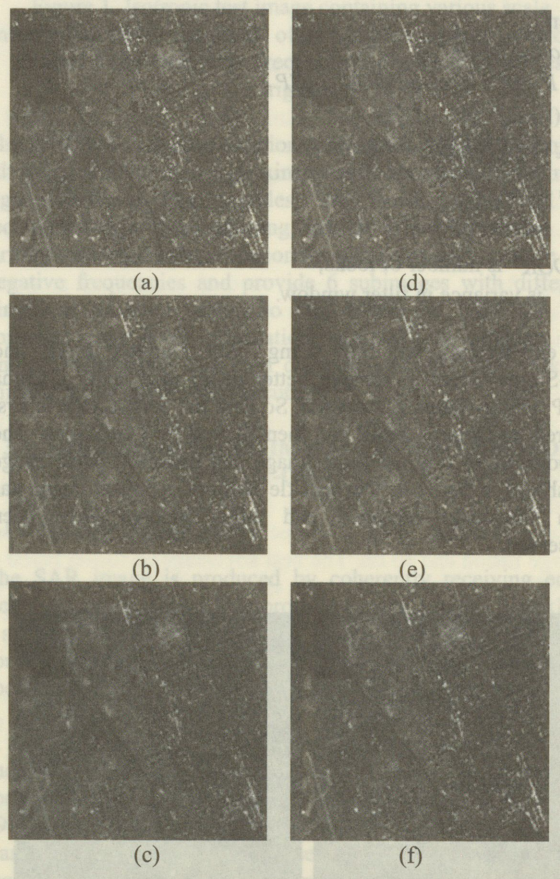

Figure 5. Proportions of images fused by: (a) 1-level DWT, (b) 2-level DWT, (c) 3-level DWT, (d) 1-level DT-CWT, (e) 2-level DT-CWT, (f) 3-level DT-CWT

In figure 5 the features in SAR image have been successfully integrated with SPOT5 image. Parts of urban area have been enhanced and spectral information has been conserved well. Figure 5 (a), (b), (c) are fused by DWT at 1, 2, 3 level respectively, and figure 5 (d), (e), (f) are fused by DT-CWT at 1, 2, 3 level respectively. We observe (a) is similar with (d) but (d) is vivider. And it is clear that (b) and (c) lost more spectral

information than (e) and (f).

Then we evaluate the performance of the fusion method using some image quality indexes. The indexes we selected are average value, standard difference, entropy, average grads and correlated coefficient. Average value can show the distribution of the image grayscale in the roughness. Standard difference and entropy can measure the information abundance in the image. Average grads shows exiguous contrast, varied texture characteristic and definition of the image. Correlated coefficient is calculated between fused image and SPOT5 multi-spectral image, which shows how much spectral information have been conserved. The statistics is shown in tab 1.

Average values in table 1 show that the mean gray level of fused image is very close to the SPOT5 image. Standard difference, entropy and average grads in table 1 show that the information insufficiency of SAR image has decreased the information insufficiency of fused images, but the fused images with DT-CWT have conserved more information than DWT at corresponding level. Correlated coefficients in table 1 show that spectral characteristics of the fused images with DT-CWT are closer to the multi-spectral SPOT5 image than DWT at corresponding level.

In a word, not only at information enhancement but also at spectral information conservation, fusion based on DT-CWT is superior to DWT.

5. CONCLUSIONS

In this paper, the dual-tree complex wavelet transform has been used to fuse SAR and optical images. According to the characteristic of SAR image, it is first denoised by Gamma MAP and Lee-Sigma filter. And then the low and high frequency parts of decomposed SAR and optical images have been fused by maximum gray value rule and maximum energy value rule respectively. Finally images fused by DT-CWT and DWT at different level have been compared in experiments. Observation results and statistics of quality indexes have shown that the fusion algorithm based on DT-CWT was better than DWT.

REFERENCES

A. Jalobeanu , L. Blanc-Féraud , J. Zerubia, 2000, Satellite image deconvolution using complex wavelet packets, Thème 3-Interaction homme-machine,images, données, connaissances Projet Ariana Rapport de recherche, pp.3955-4071

Baraldi A. , F. Parmiggiani, 1995, A refined gamma map sar speckle filter with improved geometrical adaptivity, *IEEE Transactions on Geoscience and Remote Sensing*, 33(5), pp.1245-1257

Javier Portilla , Eero P. Simoncelli, 1999, Texture modeling and synthesis using joint statistics of complex wavelet coefficients, *Proceedings of the IEEE Workshop on Statistical and Computational Theories of Vision*, Fort Collins, CO, http://www.cis.ohio-state.edu/_szhu/SCTV99.html

Jiang Han-ping, Wang Jian-min, 2000, Application of complex-valued wavelet transform to image matching, *Journal of Jiang-xi Institute of Education (Natural Science)*, 21(6), pp.29-31

Julian Magarey, Nick Kingsbury, 1998, Motion estimation using a complex_valued wavelet transform, *IEEE Trans. on Signal Processing, special issue on wavelets and filter banks*, 46(4), pp.1069-1084

Lee S. J., 1980, Digital image enhancement and noise filtering by use of local statistics, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAM1-2 (2)

Lopes A., E. Nezry, R. Touzi, H. Laur, 1993, Structure detection and statistical adaptive speckle filtering in sar images, *Int. J. Remote Sensing*, 14(9), pp.1735-1758

N. Kingsbury, 1998a, The dual-tree complex wavelet transform: A new technique for shift-invariance and directional filters, *in DSP Workshop*

Nick Kingsbury , 1998b, The dual_tree complex wavelet transform: A new efficient tool for image restoration and enhancement, *Proc. European Signal Processing Conference, EUSIPCO 98, Rhodes*, pp.319-322

Peter de Rivaz and Nick Kingsbury, 2001, Bayesian image deconvolution and denoising using complex wavelets, *Proc. IEEE Conf. on Image Processing*, Greece, Oct 8-10, paper 2639.

Serkan Hatipoglu, Sanjit K. Mitra, and Nick Kingsbury, 1999, Texture classification using dual-tree complex wavelet transform, *IEE. Image Processing and its Applications*, Conference Publication No.465

S G Mallat, 1989, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. PAMI*, 11(7), pp.674-693

Xiao Guochao, Zhu Caiying, 2001, *Radar photogrammetry*, Earthquake press, Beijing, pp.35

Images	Average Value	Standard Difference	Entropy	Average Grads	Correlated Coefficients
SAR	53.340197	26.879212	5.037007	8.630476	—
SPOT5 multi-spectral	**94.890384**	**27.411184**	**11.468343**	**12.181283**	—
Fused by 1-level DWT	97.061614	29.517138	11.524848	11.650854	0.836114
Fused by 2-level DWT	96.100397	27.015098	11.397434	9.548232	0.843378
Fused by 3-level DWT	95.309379	25.254701	11.284076	8.770403	0.833622
Fused by 1-level DT-CWT	**97.127747**	**29.676878**	**11.560681**	**12.729831**	**0.873973**
Fused by 2-level DT-CWT	**96.097643**	**27.800316**	**11.476087**	**11.813529**	**0.914931**
Fused by 3-level DT-CWT	**95.367769**	**26.061456**	**11.375349**	**11.354203**	**0.908619**

Table 1. Statistics of quality indexes of the images in figure 4 and 5