Multi-focus image fusion algorithm based on shearlets

Qiguang Miao (苗启广)[1*], Cheng Shi (石 程)[1], Pengfei Xu (许鹏飞)[1], Mei Yang (杨 眉)[1], and Yaobo Shi (史耀波)[2]

[1]*School of Computer Science, Xidian University, Xi'an 710071, China*

[2]*School of Economics and Management, Xi'an University of Technology, Xi'an 710048, China*

[*]*Corresponding author: qgmiao@163.com*

Received September 7, 2010; accepted November 24, 2010; posted online March 14, 2011

Shearlets not only possess all properties that other transforms have, but also are equipped with a rich mathematical structure similar to wavelets, which are associated to a multi-resolution analysis. Recently, shearlets have been used in image denoising, sparse image representation, and edge detection. However, its application in image fusion is still under study. In this letter, we study the feasibility of image fusion using shearlets. Fusion rules of larger high-frequency coefficients based on regional energy, regional variance, and absolute value are proposed because shearlet transform can catch detailed information in any scale and any direction. The fusion accuracy is also further improved by a region consistency check. Several different experiments are adopted to prove that fusion results based on shearlet transform can acquire better fusion quality than any other method.

OCIS codes: 100.7410, 100.0100, 100.2000.

doi: 10.3788/COL201109.041001.

Image fusion is a process that combines two or more source images to form a new, sharper, and more credible image. Using certain algorithms, image fusion produces a fused image containing the best aspects of the source images[1]. Multi-focus image fusion is an important part of fusion research, which can effectively improve the utilization of image information widely used in image sensing, computer vision, medical analysis, and so on. Capturing near and far objects in good focus using only two or three different focus settings is possible, but obtaining a completely clear image is difficult. Hence, we can acquire a series of pictures with different focus objects and fuse them into one image. Accordingly, we present an efficient algorithm for multi-focus image fusion.

Image decomposition is important to image fusion and affects information extraction quality, even the whole fusion quality. Wavelet theory has developed since the beginning of the last century. It was first applied to signal processing in the 1980s[1]. Over the past decade, wavelet theory has been recognized as having great potential in image processing applications, as well as in image fusion[2]. Wavelet transforms are more useful than Fourier transforms[3], and are more efficient in dealing with one-dimensional (1D) pointwise smooth signal. Owing to the limitations of the direction, it does not perform as well with multidimensional data. For images containing sharp transitions such as edges, wavelet transforms are not optimally efficient in representing them. Recently, a theory for multidimensional data called multi-scale geometric analysis (MGA) has been developed. Many new MGA tools have been proposed, such as ridgelet, curvelet, bandelet, contourlet, etc.[4−6], which provide higher directional sensitivity than wavelets. Shearlets, a new approach proposed in 2005, not only possess all the above properties, but also are equipped with a rich mathematical structure similar to wavelets, which are associated to a multi-resolution analysis. Shearlets form a tight frame at various scales and directions, and are optimally sparse in representing images with edges. Only certain curvelets are known to satisfy similar sparsity properties. However, the construction of curvelets is not built directly in the discrete domain and it does not provide a multi-resolution representation of the geometry. Shearlet decomposition is similar to that of contourlets; while the contourlet transform consists of an application of the Laplacian pyramid followed by directional filtering, from the shearlets, the directional filtering is obtained using a shear matrix. An important advantage of shearlet transform over contourlet transform is that there is no restriction on the number of directions for shearing. Recently, many researchers introduced shearlets into image processing[7−11]. Shearlets have been used in image denoising, sparse image representation[9,10], and edge detection[11]. Its applications in image fusion are still under exploration.

Multi-focus image fusion using shearlets is discussed in this letter. The theory of shearlets is described, and image fusion algorithm using shearlets is presented in detail. Experimental results prove the effectiveness of the shearlet-based image fusion algorithm.

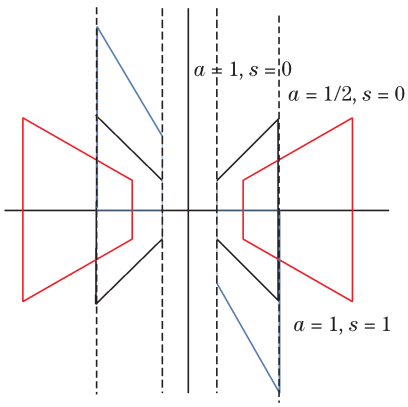


Fig. 1. Frequency support of shearlets $\psi_{j,l,k}$ for different values of a and s.

In dimension $n = 2$, affine systems with composite dilations are defined as

$$A_{\rm AS}(\psi) = \{\psi_{j,l,k}(x) = |\det A|^{j/2}\psi({\rm S}^l A^j x - k) : j, l \in \mathbb{Z}, k \in \mathbb{Z}^2\}, \quad (1)$$

where $\psi \in L^2(\mathbb{R}^2)$; A and S are both 2×2 invertible matrices, and $|\det S| = 1$. The elements of this system are called composite wavelet if $A_{\rm AS}(\psi)$ forms a tight frame for $L^2(\mathbb{R}^2)$:

$$\sum_{j,l,k} |< f, \psi_{j,l,k} >|^2 = \big|\big|f\big|\big|^2.$$

Let A denote the parabolic scaling matrix and S denote the shear matrix. For each $a > 0$ and $s \in \mathbb{R}$,

$$A = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, \quad S = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}.$$

The matrices described above have special roles in shearlet transform. The matrix A controls the "scale" of the shearlets by applying a fine dilation faction along the two axes. This ensures that frequency support of the shearlets becomes increasingly elongated at finer scales. The matrix S, on the other hand, is not expansive and only controls the orientation of the shearlets. The frequency support size of the shearlets is illustrated in Fig. 1 for some particular values of a and s.

In Ref. [8], it is assumed that $a = 4, s = 1$, $A = A_0$ is the anisotropic dilation matrix, and $S = S_0$ is the shear matrix:

$$A_0 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, S_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

For $\forall \xi = (\xi_1, \xi_2) \in \hat{\mathbb{R}}^2$, $\xi_1 \neq 0$, let $\hat{\psi}^{(0)}(\xi)$ be given by

$$\hat{\psi}^{(0)}(\xi) = \hat{\psi}^{(0)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\hat{\psi}_2(\frac{\xi_2}{\xi_1}),$$

where $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ is a wavelet and ${\rm supp}\hat{\psi}_1 \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$; $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ and ${\rm supp}\hat{\psi}_2 \subset [-1, 1]$. This implies that $\hat{\psi}^{(0)} \in C^\infty(\mathbb{R})$ and ${\rm supp}\hat{\psi}^{(0)} \subset [-\frac{1}{2}, \frac{1}{2}]^2$.

In addition, we assume that

$$\sum_{j\geq 0} |\hat{\psi}_1(2^{-2j}\omega)|^2 = 1, \quad |\omega| \geq \frac{1}{8}, \quad (2)$$

and for $\forall j \geq 0$

$$\sum_{l=-2^j}^{2^j-1} |\hat{\psi}_2(2^j\omega - l)|^2 = 1, \quad |\omega| \leq 1. \quad (3)$$

There are several examples of functions ψ_1 and ψ_2 satisfying the properties described above. Equations (2) and (3) imply that

$$\sum_{j\geq 0}\sum_{l=-2^j}^{2^j-1} |\hat{\psi}^{(0)}(\xi A_0^{-j} S_0^{-l})|^2 = \sum_{j\geq 0}\sum_{l=-2^j}^{2^j-1} |\hat{\psi}_1(2^{-2j}\xi_1)|^2 |\hat{\psi}_2(2^j\frac{\xi_2}{\xi_1} - l)|^2 = 1,$$

for any $(\xi_1, \xi_2) \in D_0$, where $D_0 = \{(\xi_1, \xi_2) \in \hat{\mathbb{R}}^2 : |\xi_1| \geq \frac{1}{8}, |\xi_2| \leq 1\}$, the functions $\{\hat{\psi}^{(0)}(\xi A_0^{-j} S_0^{-l})\}$ form a tiling of D_0. This is illustrated in Fig. 2(a). The property described above implies that the collection $\{\psi^{(0)}_{j,l,k}(x) = 2^{\frac{3j}{2}}\psi^{(0)}(S_0^l A_0^j x - k) : j \geq 0, -2^j \leq l \leq 2^j - 1, k \in \mathbb{Z}^2\}$ is a Parseval frame for $L^2(D_0)^\vee = \{f \in L^2(\mathbb{R}^2) : {\rm supp}\hat{f} \subset D_0\}$. From the conditions on the support of $\hat{\psi}_1$ and $\hat{\psi}_2$, one can easily observe that the function $\psi_{j,l,k}$ has frequency support,

$${\rm supp}\hat{\psi}^{(0)}_{j,k,l} \subset \{(\xi_1, \xi_2) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], |\frac{\xi_2}{\xi_1} + l2^{-j}| \leq 2^{-j}\}$$

That is, each element $\hat{\psi}_{j,l,k}$ is supported on a pair of trapezoids of approximate size $2^{2j} \times 2^j$, oriented along the lines of slope $l2^{-j}$ (see Fig. 2(b)).

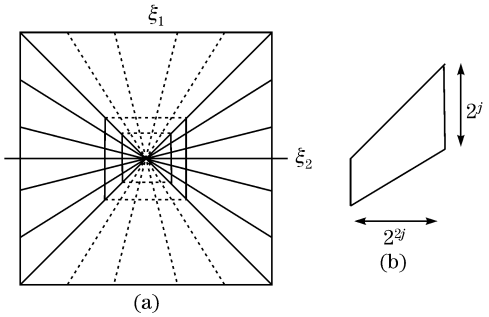


Fig. 2. Tiling of the frequency by the shearlets; (b) size of the frequency support of shearlet $\psi_{j,l,k}$.

Similarly, we can construct a Parseval frame for $L^2(D_1)^\vee$, where D_1 is the vertical cone,

$$D_1 = \{(\xi_1, \xi_2) \in \hat{\mathbb{R}}^2 : |\xi_2| \geq \frac{1}{8}, |\frac{\xi_1}{\xi_2}| \leq 1\}.$$

Let $A_1 = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$, $S_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, and $\hat{\psi}^{(1)}(\xi) = \hat{\psi}^{(1)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_2)\hat{\psi}_2(\frac{\xi_1}{\xi_2})$, where $\hat{\psi}_1$ and $\hat{\psi}_2$ are defined as Eqs. (2) and (3), respectively. Thus, the Parseval frame for $L^2(D_1)^\vee$ is as

$$\{\psi^{(1)}_{j,l,k}(x) = 2^{\frac{3j}{2}}\psi^{(1)}(S_1^l A_1^j x - k) : j \geq 0, -2^j \leq l \leq 2^j - 1, k \in \mathbb{Z}^2\}.$$

Image decomposition based on shearlet transform is composed of two parts: multi-direction decomposition and multi-scale decomposition.

1) The multi-direction part decomposes an image using

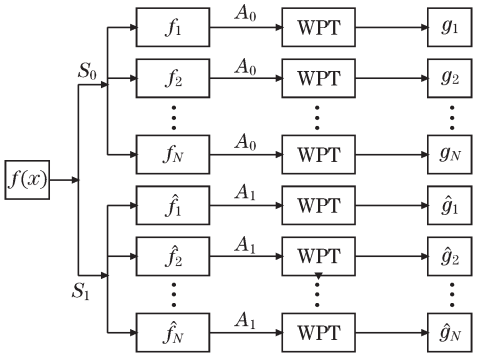


Fig. 3. Image decomposition framework with shearlets. WPT: wavelet packets transform.

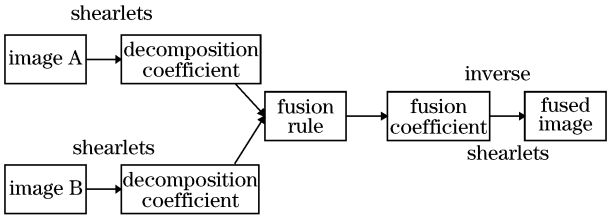


Fig. 4. Image fusion framework based on shearlets.

shear matrix S_0 or S_1.

2) The multi-scale part decomposes each direction using wavelet packets decomposition.

In step 1), if the image is decomposed only by S_0 or S_1, the number of directions is $2(l+1)+1$. If the image is decomposed by both S_0 and S_1, the number of directions is $2(l+2)+2$. The image decomposition framework with shearlets is shown in Fig. 3.

The image fusion framework based on shearlets is shown in Fig. 4. The following steps of image fusion are adopted.

1) The two images taking part in the fusion are geometrically registered to each other.

2) Transform the original images using shearlets. Both horizontal and vertical cones are adopted in this method. The number of directions is 6. Wavelet packets are then used in multi-scale decomposition with $j=5$.

3) The following fusion rules are adopted in this algorithm.

Rule 1: Fusion rule based on pixel level. Low frequency coefficients of the fused image are replaced by the average of low frequency coefficients of the two source images. High frequency coefficients are obtained by selecting the corresponding larger absolute value of high frequency coefficients.

Rule 2: Fusion rule based on regional absolute value. Low frequency coefficients of the fused image are replaced by the average of low frequency coefficients of the two sources images. Compared with the regional absolute value of high frequency coefficients of the two source images, high frequency coefficients of the fused image are obtained by the larger one. Finally, region consistency check is performed based on the fuse-decision map.

$$D_X(i,j)=\sum_{i\le M,j\le N}|Y_X(i,j)|,\quad X=\mathrm{A},\mathrm{B}, \tag{4}$$

where A and B correspond to the two source images,

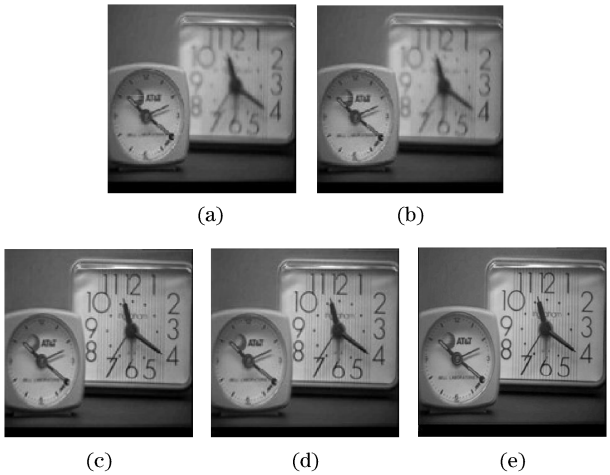


Fig. 5. Comparative experiments of fusion rules based on shearlets. (a) Focus on the left image; (b) focus on the right image; (c) Rule 1; (d) Rule 2; (e) Rule 3.

Table 1. Comparison of Image Fusion Rules

	SP	STD	EN
Rule 1	24.1629	52.9021	7.4026
Rule 2	23.4875	52.3040	7.4181
Rule 3	23.7856	52.6316	7.5108

respectively. Calculate the absolute value of high frequency coefficients in the neighborhood by Eq. (4), where $M=N=3$. The fused map is shown in

$$\mathrm{Map}(i,j)=\begin{cases}1 & D_\mathrm{A}(i,j)\ge D_\mathrm{B}(i,j)\\ 0 & D_\mathrm{A}(i,j)<D_\mathrm{B}(i,j)\end{cases}. \tag{5}$$

According to Eq. (5), if a certain pixel is to come from source image A, but with the majority of its surrounding neighbors from B, this pixel will be switched to come from B.

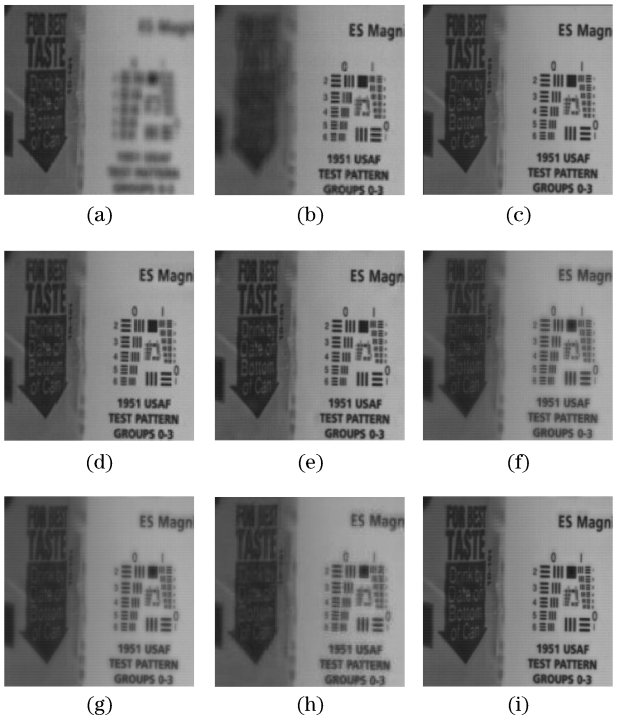
Rule 3: Fusion rule based on regional variance. Low frequency coefficients of the fused image are replaced by the average of low frequency coefficients of the two source images. Compared with the regional variance of high frequency coefficients of the two source images, high frequency coefficients of the fused image are obtained by the larger one.

$$D_X(i,j)=\frac{\sum\limits_{i\le M,j\le N}[Y_X(i,j)-\bar{Y}_X]^2}{M\times N},\quad X=\mathrm{A},\mathrm{B}. \tag{6}$$

Calculate the variance of high frequency coefficients in the neighborhood using Eq. (6), where $M=N=3$. The fused map is shown in Eq. (5). Finally, region consistency check is performed based on the fuse-decision map.

4) The fused image is obtained using the inverse shearlet transform.

Subjective visual perception gives us direct comparisons. Some objective image quality assessments are also used to evaluate the performance of the proposed approach. The following image quality metrics are used in this letter.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Fig. 6. Comparative experiments of multi-focus image fusion. (a) Focus on the left image; (b) focus on the right image; (c) original image; (d) shearlet; (e) contourlet; (f) Haar; (g) PCA; (h) Daubechies; (i) LP.

1) Entropy (EN). EN reflects the amount of information in the fused image. The larger the EN, the more information the image carries.

$$\mathrm{EN} = -\sum_{i=0}^{255} P_i \log_2 P_i, \tag{7}$$

where P_i is the ratio of the number of pixels with gray value equal to i over the total number of pixels.

2) Difference of entropy (DEN). DEN indicates the difference between fused image and original image. A smaller DEN means that there is less difference between them.

$$\mathrm{DEN} = |\mathrm{EN_f} - \mathrm{EN_o}|,$$

where $\mathrm{EN_f}$ is the entropy of the fused image and $\mathrm{EN_o}$ is the entropy of the original image.

3) Overall cross entropy (OCE). OCE can reflect the difference between the two source images and the fused image. The smaller the OCE, the better the fusion results.

$$\mathrm{OCE} = \frac{\mathrm{CE}(f_\mathrm{A}, f) + \mathrm{CE}(f_\mathrm{B}, f)}{2},$$

where f_A and f_B are the two source images, f is the fused image, and CE is the cross entropy:

$$\mathrm{CE} = \sum_{i=0}^{255} P_G(i) \log |\frac{P_G(i)}{P_f(i)}|,$$

$G =$ A or B.

4) Standard deviation (STD). STD indicates the dispersion degree between gray values and gray mean values. The larger the STD is, the more dispersed the gray level is.

$$\mathrm{STD} = \sqrt{\frac{\sum_{i=0}^{N-1}\sum_{j=1}^{M-1}\Big[(x,y)-\mu\Big]^2}{MN}},$$

where $f(x,y)$ is the pixel value of the fused image, μ and $M \times N$ are the mean value and the size of the fused image, respectively.

5) Sharpness (SP). SP can reflect the small details of the image. The larger the SP is, the better the fusion results are.

$$\mathrm{SP} = \frac{\sum_{x,y}\sqrt{\frac{[f(x,y)-f(x,y-1)]^2+[f(x,y)-f(x-1,y)]^2}{2}}}{MN}.$$

6) Mean square error (MSE). MSE indicates the dispersion degree between fused image and original image. A smaller MSE means that there is less difference between them.

$$\mathrm{MSE} = \sqrt{\frac{\sum_{i=0}^{N-1}\sum_{j=1}^{M-1}(x,y) - f_\mathrm{o}(x,y)^2}{MN}},$$

where $f_\mathrm{o}(x,y)$ is the pixel value of the original image.

7) Peak signal-to-noise ratio (PSNR). PSNR can reflect the quality of reconstruction. The larger the PSNR, the less the image distortion.

$$\mathrm{PSNR} = 10 \times \log\Big(\frac{255^2}{\mathrm{MSE}}\Big).$$

8) Mutual information (MI). MI shows the correlation of the two images. The larger the MI, the closer the information between fused image and original image.

$$\mathrm{MI}(f,o) = \frac{\mathrm{EN_f} + \mathrm{EN_o}}{\mathrm{EN_{fo}}},$$

where $\mathrm{EN_f}$, $\mathrm{EN_o}$ are the entropies of the fused image and the original image, $\mathrm{EN_{fo}}$ is the joint entropy of the two images. $\mathrm{EN_f}$, $\mathrm{EN_o}$, and $\mathrm{EN_{fo}}$ can be calculated by Eq. (6).

The fusion results for comparison of fusion rules are shown in Fig. 5 and Table 1. Figures 5(a) and (b) are the images focused on different objects; Figs. 5(c)−(e) are the fused images with three different fusion rules described above using shearlet transform. From Fig. 5 and Table 1, high sharpness, standard deviation, and entropy are achieved regardless of what fusion rule is adopted. Thus, the problem that different rules cause quite different fusion results can be avoided by shearlets, because shearlets have flexible direction features that can capture the edge information better. Fusing image using shearlets can not only get a better fused image, but also facilitate real-time image fusion by choosing lower computational complexity rules.

Different fusion methods are compared for multi-focus image fusion. Figures 6(a) and (b) are multi-focus images, Fig. 6(c) is the original image, Figs. 6(d)−(i)

Table 2. Comparison of Multi-Focus Image Fusion

	MI	DEN	OCE	SP	PSNR	MSE
Shearlet	6.7133	0.0057	0.0105	18.3345	42.5119	3.2439
Contourlet	6.0162	0.0427	0.0179	18.8710	39.3248	7.1783
Haar	6.0195	0.0180	0.0450	15.5285	31.4873	45.8468
Daubechies	6.0623	0.0176	0.0465	15.0278	31.1822	49.0835
PCA	6.3475	0.0113	0.0484	12.9532	31.1887	49.4549
LP	6.2825	0.0354	0.0179	19.4853	40.3666	5.9761

are the fused images using different methods. Fusion methods used here are shearlets, contourlets, Haar, Daubechies, principal components analysis (PCA), and Laplacian pyramid (LP). To compare fusion results under the same condition, and because regional variance is more useful for multi-focus image fusion, the fusion rule adopted is Rule 3. The fusion results are shown in Fig. 6 and Table 2. From the subjective evaluation of Fig. 6 and objective metrics of Table 2, we see that image fusion based on shearlet can better keep the detailed information and edge feature than any other method.

In conclusion, we succeed in demonstrating that shearlets are very competitive for multi-focus image fusion. As a novel MGA tool, shearlets have directionality, localization, anisotropy, and multi-scale features, and are also equipped with a rich mathematical structure similar to wavelets, which are associated to a multi-resolution analysis. Experimental results show that shearlets clearly contain more detail and smaller distortion information of the images. They are also useful for real-time fusion and multi-focus image fusion. Nonetheless, shearlets need further study, especially in its theory and applications. We will focus on other image processing methods using shearlets in our future work.

This work was jointly supported by the National Natural Science Foundation of China (Nos. 61072109 and 60702063), the Fundamental Research Funds for the Central Universities, and the Creative Project of the Science and Technology State of Xi'an under Grant (No. CXY1015(3)), and partly supported by the National Natural Science Foundation of China (Nos. 60804021 and 60802084).

References

1. S. G. Mallat, IEEE Trans. Pattern Anal. Mach. Intell. **11,** 674 (1989).
2. K. Amolins, Y. Zhang, and P. Dare, WISPRS J. Photogram. Remote Sens. **62,** 249 (2007).
3. Q. Wang, Z. Li, J. Lai, and A. He, Chin. Opt. Lett. **8,** 278 (2010).
4. M. N. Do and M. Vetterli, IEEE Trans. Image Process. **14,** 2091 (2005).
5. L. Yang, B. L. Guo, and W. Ni, Neurocomputing **72,** 203 (2008).
6. Q. Li, Y. Zeng, X. Peng, and K. Yang, Chin. Opt. Lett. **8,** 577 (2010).
7. G. Kutyniok and D. Labate, Trans. Am. Math. Soc. **361,** 2719 (2009).
8. G. Kutyniok and D. Labate, Journal of Wavelet Theory and Applications **1,** 1 (2007).
9. G. R. Easley, D. Labate, and W.-Q Lim, Appl. Comput. Harmon. Anal. **25,** 25 (2008).
10. K. Guo and D. Labate, SIAM J. Math. Anal. **39,** 298 (2007).
11. K. Guo, D. Labate, and W.-Q Lim, Appl. Comput. Harmon. Anal. **27,** 24 (2009).