Fusion of infrared and visible images based on focus measure operators in the curvelet domain

Shao Zhenfeng,[2,*] Liu Jun,[1] and Cheng Qimin[3]

[1]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

[2]State Key Laboratory for Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China

[3]The Department of Electronics and Information Engineering, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China

*Corresponding author: shaozhenfeng@163.com

Received 29 November 2011; revised 7 February 2012; accepted 7 February 2012; posted 7 February 2012 (Doc. ID 159126); published 11 April 2012

Aiming at the differences of physical characteristics between infrared sensors and visible ones, we introduce the focus measure operators into the curvelet domain in order to propose a novel image fusion method. First, the fast discrete curvelet transform is performed on the original images to obtain the coefficient subbands in different scales and various directions, and the focus measure values are calculated in each coefficient subband. Then, the local variance weighted strategy is employed to the low-frequency coefficient subbands for the purpose of maintaining the low-frequency information of the infrared image and adding the low-frequency features of the visible image to the fused image; meanwhile, the fourth-order correlation coefficient match strategy is performed to the high-frequency coefficient subbands to select the suitable high-frequency information. Finally, the fused image can be obtained through the inverse curvelet transform. The practical experiments indicate that the presented method can integrate more useful information from the original images, and the fusion performance is proved to be much better than the traditional methods based on the wavelet, curvelet, and pyramids.

OCIS codes: 100.0100, 100.2980.

1. Introduction

The infrared and visible sensors are the two most commonly used ones. They work in different wave bands and can provide complementary image information. The infrared sensor records the infrared radiation information through the acquisition of infrared radiation emitted by ground targets. Because the infrared sensor identifies the targets by detecting the thermal radiation difference between target and background, it has a special ability to recognize camouflage, while it has less sensitivity to the luminance change and has lower contrast. In contrast, the visible sensor is more sensitive to the reflection of a scene with a low noise level and a high definition, so it can provide more detail information about the scene. To sum up, the fusion of these two types of images is useful to integrate good target characteristics in infrared images and clear scene information in visible images [1,2].

So far, the multiscale analysis methods represented by wavelet transform and pyramid decomposition have achieved great success in fusion of infrared and visible images [3–5]. It is generally believed that the fusion effectiveness of wavelet transform is superior to that of pyramid decomposition [6]. However, despite many advantages, the wavelet transform still cannot accurately express the edge direction of an image. Since the wavelet is in lack of

1559-128X/12/121910-12$15.00/0

translation invariance, there is a block effect in the image edges. To solve this problem, Candès and Donoho proposed the curvelet transform based on the best approximation of curve and image sparse representation, which had been successfully used in image fusion [7,8]. Being compared with the wavelet transform, the advantage of the curvelet transform is the high anisotropy of all characteristics in detail scale that can better approximate curves and describe the details characteristics and directional information such as edges and textures of an image.

When the multiscale analysis method is applied to image fusion, the fusion rules are essential to the fusion methods. Traditional rules such as average or maximum modulus will lead to the quality descent of fused image. Focus measure operators derive from the multifocus image fusion domain and are used to measure the rate of pixel change. The values of focus measure operators will be bigger in the better-focused areas of an image [9]. Considering that the visible image is rich with details, the targets are obvious in the infrared image, and the coefficient subbands of an image in the curvelet domain can be treated as an image with amplitude change. Hence, it is reasonable to choose the coefficients in the curvelet domain by using focus measure operators. In this paper, the focus measure operators are introduced into the curvelet domain, and an infrared and visible image fusion method is proposed. We use the local variance weighted strategy to fuse the low-frequency information and the fourth-order correlation coefficient match strategy to obtain the high-frequency coefficients. The experimental results indicate that the proposed method can extract more useful information from the original images and show obvious advantages in visual effect and performance in contrast to the traditional methods.

The rest of this paper is organized in the following sequence. In Sections 2 and 3, we review theories of curvelet transform and focus measure operators. In Section 4, the proposed infrared and visible image fusion method is introduced. Some experimental results are shown to demonstrate the validity of the method in Section 5. Concluding remarks are summarized in the final section.

2. Curvelet Transform

Researchers have achieved great success in the theory and applications of the curvelet transform over a decade, and two generations of curvelet have been developed. The principle of the first-generation curvelet is to approximate the curve to straight lines in each block through small enough blocking, and to analyze its properties using the local ridgelet transform. In 2004, Candès and Donoho proposed the second-generation curvelet transform [10], and realized its fast algorithm in 2005 [11]. In contrast to the first generation, the second generation reduces the variable number from 7 to 3, has a simpler structure, decreases the data redundancy greatly, and is easier to understand and realize. The second generations of curvelet transform and its realization have nothing to do with the ridgelet transform. The only thing they are related to is the abstract mathematical meaning of compact support and framework and so on. In this section, we make a short review on the theory of curvelet transform.

A. Continuous Curvelet Transform

The continuous curvelet transform achieves the sparse representation of signals by using the inner product of basis function and signals with the expression as follows:

$$c(j,l,k) = \langle f, \varphi_{j,l,k} \rangle, \tag{1}$$

where $\varphi_{j,l,k}$ is the curvelet function and j, l, k denote the scale, direction, and position respectively.

We define the radial window function as $W(r)$, $r \in (\frac{1}{2}, 1)$and angle window function as $V(t)$, $t \in [-1, 1]$. They all meet the following requirements:

$$\sum_{j=-\infty}^{\infty} W^2(2^j r) = 1, \qquad r \in \left(\frac{3}{4}, \frac{3}{2}\right), \tag{2}$$

$$\sum_{l=-\infty}^{\infty} V^2(t-l) = 1, \qquad t \in \left(-\frac{1}{2}, \frac{1}{2}\right). \tag{3}$$

For each $j \geq j_0$, the frequent window U_j is defined as:

$$U_j(r,\theta) = 2^{-3j/4} W(2^{-j} r) V\left(\frac{2^{[j/2]}\theta}{2\pi}\right), \tag{4}$$

where $[j/2]$ is the integer part of $j/2$ and the support interval of U_j is a wedge area that is restricted by the support intervals W and V and has a property of anisotropy. The curvelet with the scale 2^{-j}, direction θ_l, and translation (k_1,k_2) is:

$$\varphi_{j,l,k}(x) = \varphi_j(R_{\theta_l}(x - x_k^{(j,l)})), \tag{5}$$

where $x_k^{(j,l)} = R_{\theta_l}^{-1}(2^{-j}k_1, 2^{-j/2}k_2)$ and R_{θ_l} is rotated by θ_l; thus the curvelet transform can be expressed as:

$$\begin{aligned} c(j,l,k) := \langle f, \varphi_{j,l,k} \rangle &= \int_{R^2} f(x)\overline{\varphi_{j,l,k}(x)}\mathrm{d}x \\ &= \frac{1}{2\pi^2}\int \hat{f}(\omega)\overline{\varphi_{j,l,k}(\omega)}\mathrm{d}\omega \\ &= \frac{1}{2\pi^2}\int \hat{f}(\omega) U_j(R_{\theta_l}\omega) e^{j\langle x_k^{(j,l)}, \omega\rangle}\mathrm{d}\omega. \end{aligned} \tag{6}$$

B. Discrete Curvelet Transform

The discrete curvelet transform (DCT), with the input of $f[t_1, t_2](0 \leq t_1, t_2 < n)$ in Cartesian coordinate system, can be expressed as:

$$C^D(j,l,k) := \sum_{0\le t_1,t_2<n} f[t_1,t_2]\overline{\varphi^D_{j,l,k}[t_1,t_2]}. \tag{7}$$

Band-pass function $\psi(\omega_1) = \sqrt{\varphi(\omega_1/2)^2 - \varphi(\omega_1)^2}$ and $\psi_j(\omega_1) = \psi(2^{-j}\omega_1)$ are used to realize the multiscale segmentation. For each $\omega \in (\omega_1, \omega_2)$, $\omega_1 > 0$, we have

$$V_j(S_{\theta_l}\omega) = V(2^{\lfloor j/2 \rfloor}\omega_2/\omega_1 - l), \tag{8}$$

where S_{θ_l} is a share matrix with the following expression:

$$S_{\theta_l} = \begin{bmatrix} 1 & 0 \\ -\tan\theta_l & 1 \end{bmatrix}. \tag{9}$$

If we define $\tilde{U}_j(\omega) := \psi_j(\omega_1)V_j(\omega)$, then for each $\theta_l \in [-\frac{\pi}{4}, \frac{\pi}{4})$, we have:

$$\tilde{U}_{j,l}(\omega) := \psi_j(\omega_1)V_j(S_{\theta_l}\omega) = \tilde{U}_j(S_{\theta_l}\omega). \tag{10}$$

Figure 1 is the schematic diagram of the discrete curvelet transform. The direction variance is not introduced into the coarse scale level, so similar to the wavelet transform, the curvelet here has the property of isotropy; while in scale j level, the frequency band is divided into $2^{\lfloor j/2 \rfloor}$ wedge areas by the same interval and slope. $\lfloor j/2 \rfloor$ is the integer of $j/2$. Each wedge area has an anisotropy feature that is illustrated in Fig. 1. The shaded part represents the support interval of curvelet function in a certain scale and direction.

C. Fast Discrete Curvelet Transform

Candès and Donoho proposed two fast discrete curvelet transform (FDCT) algorithms, which are unequally spaced fast Fourier transforms (USFFT) algorithm and wrapping-based transform (Wrap) algorithm respectively. The former is based on the fast Fourier transform, while the principle of the latter is surrounding with origin. From the perspectives of complexity and time-consumption, the Wrap algorithm is simpler than the USFFT one. So the proposed method in this paper is based on the Wrap algorithm that is described in detail as follows. For more details, see [11].

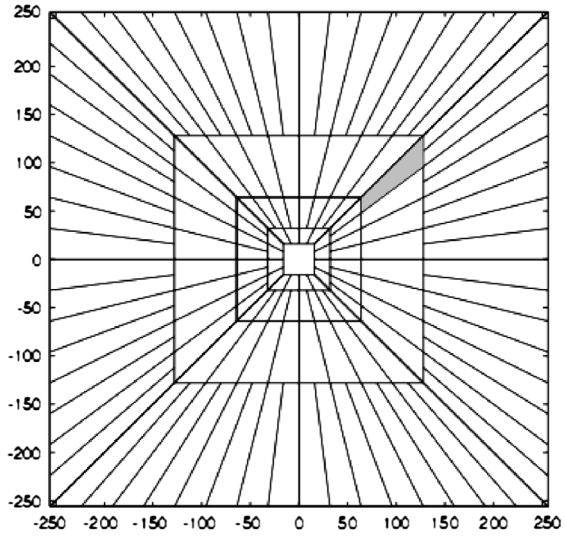

Fig. 1. Schematic diagram of the discrete curvelet transform [11].

(1) Apply the 2D FFT transform to a 2D function in Cartesian coordinates to obtain the Fourier coefficients:

$$\hat{f}[n_1,n_2], \qquad -\frac{n}{2} \le n_1, \qquad n_2 \le \frac{n}{2}(j,l). \tag{11}$$

(2) For each scale j and angle l, resample $\hat{f}[t_1,t_2]$ to obtain the resampled value:

$$\hat{f}[n_1, n_2 - n_1\tan\theta_l], \qquad (n_1,n_2) \in P_j. \tag{12}$$

(3) Multiply the resampled $\hat{f}$ and window function $\tilde{U}_j$ to make:

$$\hat{f}[n_1,n_2] = \hat{f}[n_1, n_2 - n_1\tan\theta_l]\tilde{U}_j[n_1,n_2]. \tag{13}$$

(4) Localize $\hat{f}[n_1,n_2]$ by wrapping around the origin.

(5) Apply inverse 2D FFT to $\hat{f}_{j,l}$; then we obtain the discrete curvelet coefficient sets $C^D(j,l,k)$.

3. Focus Measure Operators

So far, many focus measure operators (FMOs) have been proposed and applied in multifocus image fusion and other fields successfully [1215]. A focus measure operator must satisfy the following requirements [16].

(1) It must respond to high-frequency variations in image intensity.
(2) It must be independent of the image content.
(3) It has one and only one maximum value.
(4) It has minimal computation complexity.

In this paper, the focus measure operators are introduced into infrared and visible image fusion in the curvelet domain. We assume $f(x,y)$ represents the gray value of the pixel with the coordinate (x,y). The focus measure operators used in this paper are described below:

A. Energy of Gradient

The energy of gradient (EOG) of an image is computed by the following formula:

$$\mathrm{EOG} = \sum_x \sum_y (f_x^2 + f_y^2), \tag{14}$$

where $f_x = f(x+1,y) - f(x,y)$ and $f_y = f(x,y+1) - f(x,y)$ represent the horizontal and vertical gradient respectively.

B. Tenenbaum's Algorithm

Tenenbaum proposes a focus measure method named Tenengrad (TEN), which describes the gradient energy by employing the Sobel operator.

$$\text{Tenengrad} = \sum_{x=2}^{M-1}\sum_{y-2}^{N-1}[\nabla S(x,y)]^2, \tag{15}$$

where $\nabla S(x,y) = [\nabla S_x(x,y)^2 + \nabla S_y(x,y)^2]^{1/2}$, T is the threshold, and $\nabla S_x(x,y)$ and $\nabla S_y(x,y)$ are the horizontal and vertical Sobel gradient value respectively:

$$\begin{aligned}\nabla S_x(x,y) = \{&f(x+1,y-1) - f(x-1,y-1)\\ &+ 2f(x+1,y) - 2f(x-1,y)\\ &+ f(x+1,y+1) - f(x-1,y+1)\},\end{aligned} \tag{16}$$

$$\begin{aligned}\nabla S_y(x,y) = \{&f(x-1,y+1) - f(x-1,y-1)\\ &+ 2f(x,y+1) - 2f(x,y-1)\\ &+ f(x+1,y+1) - f(x+1,y-1)\}.\end{aligned} \tag{17}$$

C. Energy of Laplacian of Image

The energy of Laplacian (EOL) is used for analyzing the high spatial frequencies and the border sharpness of an image. It is described as follows:

$$\text{EOL} = \sum_x\sum_y (f_{xx} + f_{yy})^2, \tag{18}$$

where

$$\begin{aligned}f_{xx} + f_{yy} = &20f(x,y) - f(x-1,y-1) - 4f(x-1,y)\\ &- f(x-1,y+1) - 4f(x,y-1)\\ &- 4f(x,y+1) - f(x+1,y-1)\\ &- 4f(x+1,y) - f(x+1,y+1).\end{aligned} \tag{19}$$

D. Sum-Modified Laplacian

Nayar and Nakagawa [17] noted that the opposite signs will appear in the second derivatives in the x- and y-directions of the Laplacian, which tend to cancel each other, so a modified Laplacian (ML) is proposed by the following expression:

$$\begin{aligned}\nabla^2_{\text{ML}} f(x,y) = &|2f(x,y) - f(x-\text{step},y) - f(x+\text{step},y)|\\ &+ |2f(x,y) - f(x,y-\text{step})\\ &- f(x,y+\text{step})|,\end{aligned} \tag{20}$$

where step is a variable spacing. In this paper, step always equals 1. Hence, the sum-modified Laplacian (SML) is:

$$\text{SML} = \sum_{i=x-N}^{i=x+N}\sum_{j=y-N}^{j=y+N} \nabla^2_{\text{ML}} f(i,j), \tag{21}$$

where N is the window size used to compute the SML focus measure.

4. Proposed Fusion Method

The infrared and visible image fusion method proposed in this paper shown in Fig. 2 mainly consists of these steps: Firstly, the FDCT is applied to the infrared and visible images to obtain the corresponding low- and high-frequency coefficients in different subbands where the values of focus measure operators are calculated; then, according to the characteristics that the low frequency energies are concentrated in infrared images, we fuse the low-frequency coefficients by using local variance weighted strategy. On the basis of the characteristics that the visible images are rich of high-frequency details, the fusion of high-frequency coefficients is implemented by employing the fourth order correlation coefficient (FOCC) match strategy. Finally, the fused image is obtained by the inverse FDCT.

In this paper, the fusion rules of low- and high-frequency coefficients are the key points that are described in the following subsections.

A. Low-Frequency Coefficients Fusion Rules

The low-frequency coefficients are the reflection of energy distribution of images. According to the imaging principle of infrared and visible sensors, it is known that the former is mainly based on the object's thermal radiation characteristics, while the latter is based on spectral reflectance properties of the object. Thus, the infrared images well reflect the existence of targets, while the scene information is reflected in visible images. However, the gray value distributions

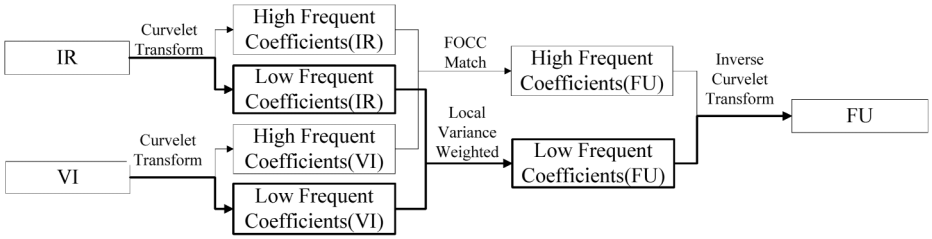

Fig. 2. Fusion scheme of infrared and visible images based on focus measure operators in the curvelet domain.

of the same target are different, or even with completely opposite polarity in these two kinds of images. Therefore, not taking into account the physical properties of the sensors, the average weighted strategy used in fusing the low-frequency coefficients in traditional methods will result in contrast reduction of the fused images and the loss of useful information.

Considering the above situation, we first calculate the focus measure values of the low-frequency coefficients:

$$C_{\text{inf}}^{l} = \text{FOM} \otimes C_{\text{inf}}^{\text{low}}, \tag{22}$$

$$C_{\text{vis}}^{l} = \text{FOM} \otimes C_{\text{vis}}^{\text{low}}, \tag{23}$$

where FOM denotes one of the above focus measure operators, $\otimes$ is the calculating processing, $C_{\text{inf}}^{\text{low}}$ and $C_{\text{vis}}^{\text{low}}$ are the low-frequency coefficients of infrared and visible images respectively, and C_{inf}^{l} and C_{vis}^{l} are the corresponding low-frequency coefficients.

Then the local variance weighted strategy is employed to obtain the low-frequency coefficients of the fused image:

$$C_{\text{fus}}^{l}(x,y) = C_{\text{inf}}^{l}(x,y) + \frac{\sigma(W_{\text{vis}})}{\sigma(W_{\text{vis}}) + \sigma(W_{\text{inf}})} \times [C_{\text{vis}}^{l}(x,y) - \min(\bar{W}_{\text{vis}}, \bar{W}_{\text{inf}})], \tag{24}$$

where C_{inf}^{l}, C_{vis}^{l}, and C_{fus}^{l} represent the low-frequency coefficient subbands of the infrared, visible, and fused images, W is the size of local window, σ denotes the values variance of focus measure operators, $\bar{W}$ is the mean of low-frequency coefficients, and min $(\bar{W}_{\text{vis}}, \bar{W}_{\text{inf}})$ is the minimum mean value of low-frequency coefficients of infrared and visible. These variables are all calculated in the local window.

Equation (24) can be simplified to

$$C_{\text{fus}}^{l}(x,y) = C_{\text{inf}}^{l}(x,y) + w \times [C_{\text{vis}}^{l}(x,y) - C]. \tag{25}$$

In Eq. (25), $C = \min(\bar{W}_{\text{vis}}, \bar{W}_{\text{inf}})$ denotes the common low-frequency characteristics of the infrared and visible images, while $C_{\text{vis}}^{l}(x,y) - C$ represents the particular low-frequency characteristics of the visible image. $w = \frac{\sigma(W_{\text{vis}})}{\sigma(W_{\text{vis}}) + \sigma(W_{\text{inf}})}$ is the weight parameter used to add the particular low-frequency characteristics of the visible image into the infrared image. σ is the variance that measures the energy distribution difference. After the calculation by using focus measure operators, the variance is more suitable to represent the energy distribution of low-frequency coefficients. From Eq. (25) we can see that the low-frequency coefficients of the fused image consist of the low-frequency coefficients of the infrared image and the particular low-frequency coefficients of the visible image. Thus, the fused image can not only maintain the whole information of the infrared image, but also efficiently use the particular information of the visible image to achieve the goal of maintaining the low-frequency information of original images.

B. High-Frequency Coefficients Fusion Rules

The high-frequency coefficients of the image reflect the distribution of detail characteristics. The visible images are rich in high-frequency details information, while in contrast, the infrared images are lack of details. Therefore, for all the high-frequency coefficients subbands of infrared and visible images, we first calculate the values of focus measure operators in the current subband:

$$C_{\text{inf}}^{h} = \text{FOM} \otimes C_{\text{inf}}^{\text{high}}(k), \tag{26}$$

$$C_{\text{vis}}^{h} = \text{FOM} \otimes C_{\text{vis}}^{\text{high}}(k), \tag{27}$$

where $C_{\text{inf}}^{\text{high}}(k)$ and $C_{\text{vis}}^{\text{high}}(k)$ are kth coefficient subbands of infrared and visible images, C_{inf}^{h} and C_{vis}^{h} denote the corresponding coefficients.

Then the match of the values is calculated by using the local fourth order correlation coefficients [18] to choose the proper high-frequency coefficients. The FOCC is mathematically defined as

$$\text{FOCC}_{A,B} = \frac{1}{M \times N} \times \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} (A(i,j) - \mu_A)^2 (B(i,j) - \mu_B)^2}{\sqrt{\left(\sum_{i=1}^{M} \sum_{j=1}^{N} (A(i,j) - \mu_A)^4\right)\left(\sum_{i=1}^{M} \sum_{j=1}^{N} (B(i,j) - \mu_B)^4\right)}}, \tag{28}$$

where A and B are two vectors; in the fusion procedure, they are used to denote the values of C_{inf}^{h} and C_{vis}^{h} in a certain local window respectively. μ_A and μ_B are the mean values of A and B respectively. The window size is $M \times N$. The FOCC value denotes the match value of the high-frequency coefficients of the infrared and visible images in a particular local area. When choosing the proper high-frequency coefficients, we use a threshold T; if $\text{FOCC}(x,y) < T$, we need maintain the information of both original images at the same time for the reasons that their match is low. In this situation, the high-frequency coefficients of the fused image are defined as

$$C^h_{\mathrm{fus}}(x,y) = C^h_{\mathrm{inf}}(x,y) + C^h_{\mathrm{vis}}(x,y). \quad (29)$$

When $\mathrm{FOCC}(x,y) \geq T$, their match is high. In order to decrease the information redundancy, we maintain only the high-frequency coefficients that have the bigger value. In this situation, the high-frequency coefficients of the fused image are defined as

$$C^h_{\mathrm{fus}}(x,y) = \begin{cases} C^h_{\mathrm{inf}}(x,y), & \text{if } |C^h_{\mathrm{inf}}(x,y)| > |C^h_{\mathrm{vis}}(x,y)| \\ C^h_{\mathrm{vis}}(x,y), & \text{otherwise} \end{cases}. \quad (30)$$

5. Experiment and Analysis

To test the efficiency of the proposed method, the infrared and visible images are fused, and the results are compared with those of traditional methods. We have tested many images, which are available from http://www.imagefusion.org, and drawn the same conclusion. Here we provide the results of two typical fusion experiments. For comparison purposes, the classical curvelet, gradient pyramid (GP) [19], and the classical wavelet methods are also used to fuse the same images. In the classical curvelet and wavelet methods, we get the low-frequency coefficients of the fused image by averaging those of the original images, and obtain the high-frequency coefficients of the fused image by choosing the maximum absolute values.

Considering that the objects in the infrared image should be preserved as much as possible, we use an object error histogram to evaluate how much information is transformed from the original infrared image. The object error image is a subset of difference image between the fused image and the original infrared image, which mainly contains the objects and is obtained by subtracting object regions from the fused image. The histogram of the object error image denotes the difference distribution of the object regions in the procedure of image fusion. Since in some pixels the gray value of the fused image is lower than that of the infrared image, we normalize the x-coordinate of the object error histogram from [0,255] to [−128,128]. Therefore, for a well-fused image, the value distribution of the object error histogram will concentrate near zero. The values less than zero mean that the gray values of the fused image are lower than those of the original infrared image, and vice versa. Hence, we can evaluate the performance of the fused images by checking the object error histogram.

It should be pointed out that we calculate the low- and high-frequency coefficients in a local window. In the following experiments, the window size is set to be 3. The reason is that with the increase of window size, the amounts of some calculations, such as variance and FOCC, will also increase.

A. Fusion Results Using Different Methods

The size of images in the first experiment is 360 × 270 pixels. The decomposition level of curvelet transform is 6, and the threshold of FOCC is set to be 0.9. Figure 3 illustrates the source images and the fusion results obtained by different methods. Figures 3(a) and 3(b) are the infrared and visible images respectively. In Fig. 3(a), the region labeled with white frame is the object region used to calculate the object error histogram. As shown in Fig. 3, the object (person) is obvious in the infrared image, while the environment scene details are clear in the visible image. In all fused images, we can see that the person object is as clear as the original infrared image in the results of the proposed method, while in the other methods, they are all dim. The same conclusion can be drawn from the object error histogram in Fig. 4. Since most histogram values are less than zero, the objects in the GP and wavelet-based results are darker than those of the infrared image. The curvelet-based result is better than those of GP and wavelet-based, which can be seen from the fused images and the object error histogram simultaneously. The results of the proposed methods are visually better than those of the above methods, as the object is almost preserved entirely.

In the second experiment, the image size is 505 × 510 pixels. The decomposition level of curvelet transform is 6, and the threshold of FOCC is set to 0.9. The fusion results are shown in Fig. 5, and the corresponding object error histograms are shown in Fig. 6. Since there are many objects in the source infrared images, such as persons and boats, we choose only persons, labeled in black frame, to calculate the object error histogram. Figures 5(a) and 5(b) are the source infrared and visible images respectively. From Fig. 5, the objects (persons, boats, etc.) are preserved well by the proposed method, while are weakened and low contrast in curvelet-based, wavelet-based, and GP method, and in particular, the sawtooth edges are obvious in the GP method. From the object error histogram shown in Fig. 6, we can see that almost all the histogram values are near zero, indicating that the proposed methods are much better than the other methods. For the proposed method, the low-frequency information of the infrared image is wholly preserved, and the high-frequency coefficients are chosen by the FOCC match rule. Hence, we can get the fused image with complete targets and clear background information.

Besides visual observation, for further comparison, we also use the objective assessment indices, such as entropy (EN), standard deviation (SD), definition (DE), spatial frequency (SF), mutual information (MI) [20], and Universal Image Quality Index (UIQI) [21] to evaluate the fusion results of all methods. EN reflects the richness of information in an image. The greater the EN is, the larger amount information the image contains. SD, DE, and SF reflect the details contrast of an image. A greater value means richer details. MI is a reflection that measures how much information the fused image maintains from the original images. UIQI shows the similarity between the fused and original images from the

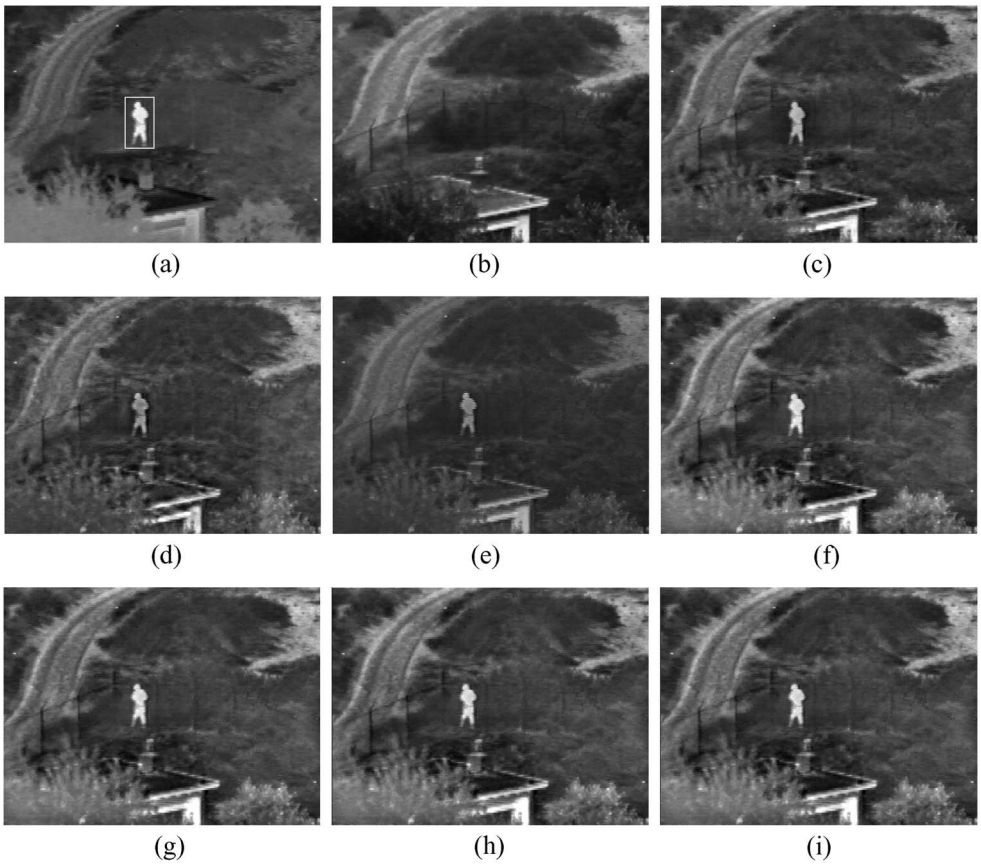

Fig. 3. Source images and results of the first experiment: (a) infrared image, (b) visible image, (c) curvelet transform, (d) gradient pyramid, (e) wavelet transform, (f) curvelet and EOG, (g) curvelet and EOL, (h) curvelet and SML, and (i) curvelet and TEN.

aspects of loss of correlation, luminance distortion, and contrast distortion. In this paper, MI is the sum value of the MI between fused and infrared image and MI between fused and visible image, while UIQI is the average value. Table 1 shows the objective evaluation results of the first experiment.

In the above table, the bold numbers are the maximum value in a column. Figure 7 gives the normalized objective evaluation results of the second experiment. In each index, the maximum value divided by each value is the normalized value. The x-coordinate denotes all the fusion methods, where the EOG represents the EOG operator in curvelet domain proposed in this paper, similar to the other operators. The y-coordinate denotes the normalized values of all the objective indices.

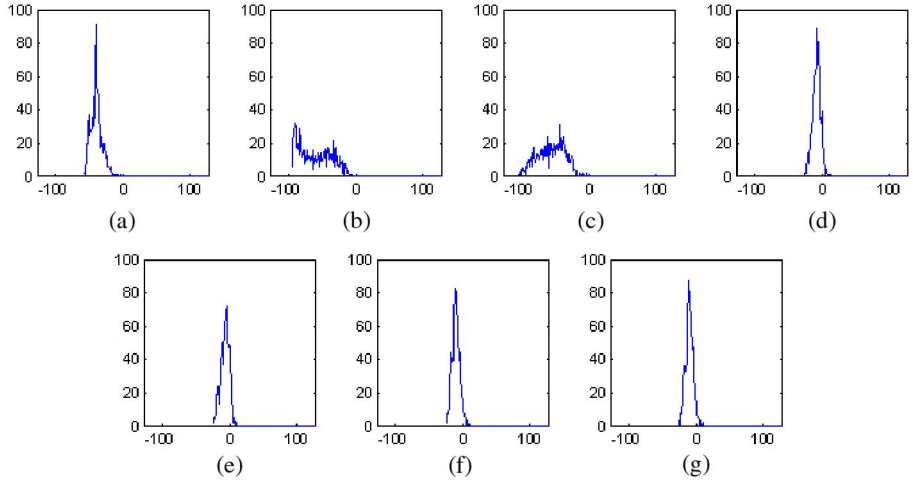

Fig. 4. (Color online) Object error histogram of the first experiment: (a) curvelet transform, (b) gradient pyramid, (c) wavelet transform, (d) curvelet and EOG, (e) curvelet and EOL, (f) curvelet and SML, and (g) curvelet and TEN.

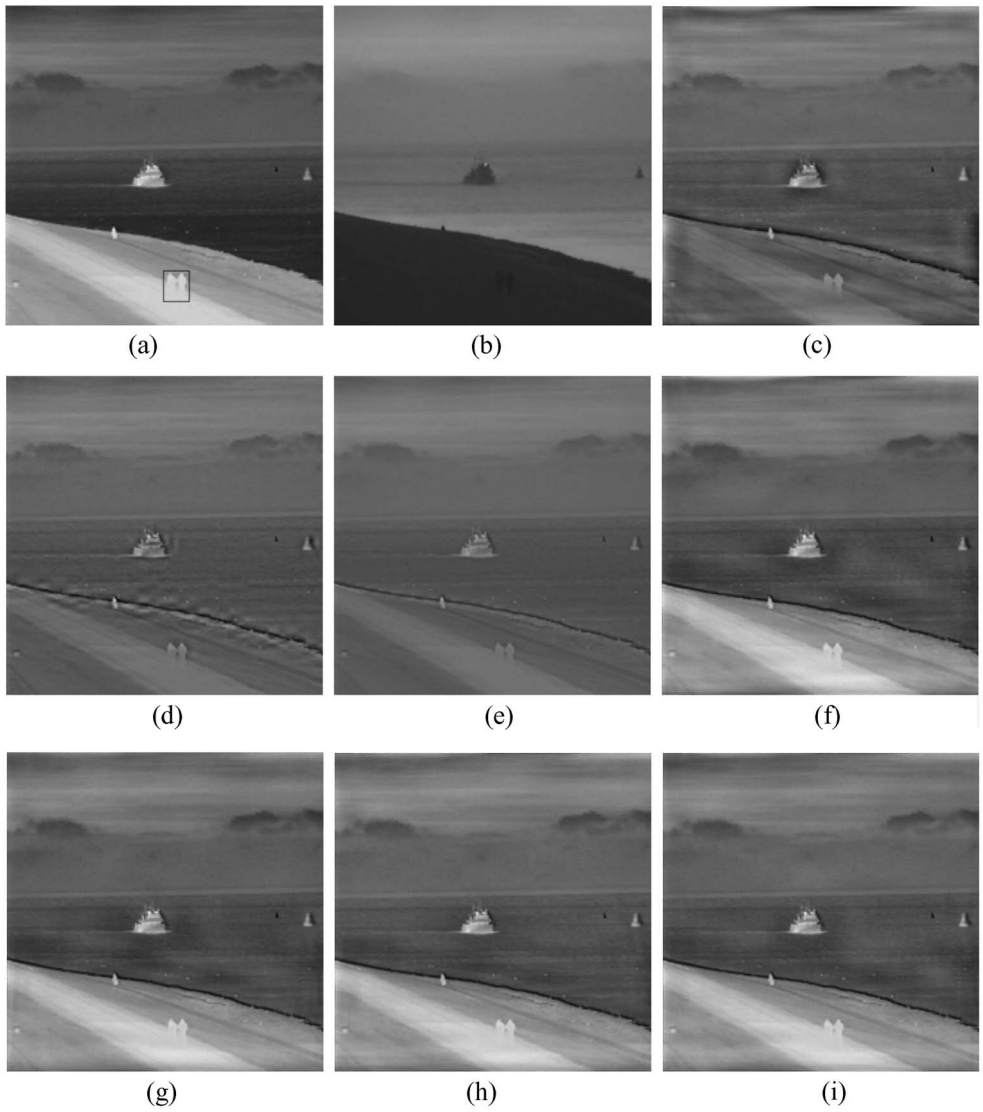

Fig. 5. Source images and results of the second experiment: (a) infrared image, (b) visible image, (c) curvelet transform, (d) gradient pyramid, (e) wavelet transform, (f) curvelet and EOG, (g) curvelet and EOL, (h) curvelet and SML, and (i) curvelet and TEN.

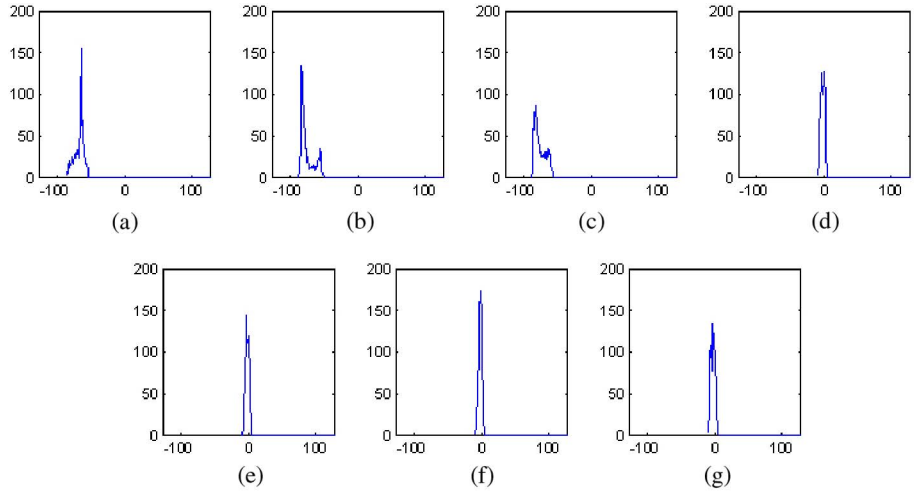

Fig. 6. (Color online) Object error histogram of the first experiment: (a) curvelet transform, (b) gradient pyramid, (c) wavelet transform, (d) curvelet and EOG, (e) curvelet and EOL, (f) curvelet and SML, and (g) curvelet and TEN.

Table 1. Objective Evaluation Results of the First Experiment

	EN	SD	DE	SF	MI	UIQI
Curvelet + EOG	7.1091	36.9891	7.6006	14.3388	1.6459	0.7865
Curvelet + EOL	7.0822	36.3541	7.5584	14.2616	1.6236	0.7750
Curvelet + SML	7.0556	35.5500	7.5601	14.2768	1.5866	0.7639
Curvelet + TEN	7.0937	36.4776	7.6011	14.3235	1.6402	0.7782
Curvelet	6.7757	31.1534	7.0879	13.4836	1.3811	0.7433
GP	6.3170	23.8827	5.6458	10.7133	1.5734	0.6881
Wavelet	6.8001	30.8361	7.3435	13.8123	1.4250	0.7178

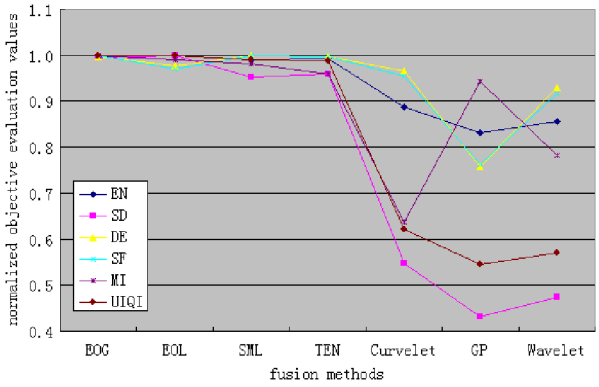

Fig. 7. (Color online) Object evaluation results of the second experiment.

According to the objective evaluation results in Table 1 and Fig. 7, the quality of the fused images produced by the proposed methods are close to each other, and are obviously better than the other three traditional methods, which are consistent with the subjective evaluation. Thus we conclude that the methods proposed in this paper can extract more useful information from the original images to inject into the fused images, which are with more outstanding performances.

B. Decomposition Level of Curvelet Transform

In order to determine how the decomposition level of curvelet transform affects the fusion results, we fused the infrared and visible images by increasing

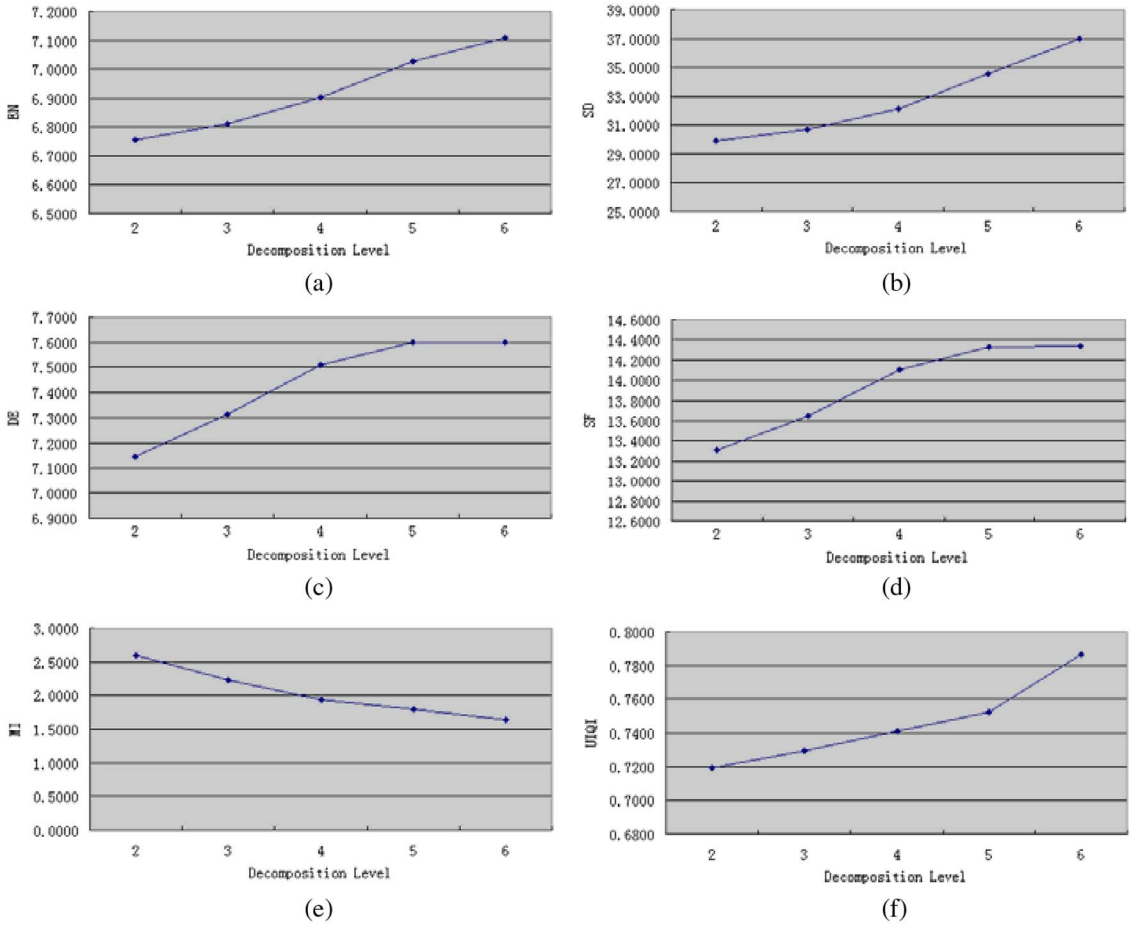

Fig. 8. (Color online) Influence of decomposition level of curvelet transform.

the decomposition level from 2 to 6. For an image with width w and height h, the maximum level of decomposition is $\text{ceil}(\log 2(\min(w,h)-3))$, where the $\text{ceil}(x)$ function is the minimum integer that is bigger than x. The fusion method is curvelet with EOG operator and the threshold is set to be 0.9. The objective assessment of the fused images with different decomposition levels are shown in Fig. 8.

Generally speaking, the assessment values increase with the increase of the decomposition levels except MI. When the decomposition level increases, the size of low-frequency band becomes smaller, and then the effect of low-frequency coefficients is lower, so more and more high-frequency information is added into the fused image. That is why the assessments such as EN, SD, DE, and SF increase. While for the MI index, with the increase of decomposition level, more and more details information is extracted and injected into the infrared image; as a result, the fused images are more and more similar to the visible image and dissimilar to the infrared image. So the MI of the fused image and the visible image will increase, and the MI of the fused image and the infrared image will decrease. However, the sum of these two MI values decreases because the latter MI value decreases more rapidly than the former one.

C. Threshold of FOCC

When applying the high-frequency coefficients fusion rules, we use a threshold to determine what coefficients should be chosen. In this experiment, we increase the threshold from 0.8 to 1 to investigate its effect on the fusion results. The fusion method is curvelet with EOG operator, and the decomposition level of curvelet transform is set to 6. Analogously, all the fused images with different thresholds are assessed using objective quality indices, and the results are shown in Fig. 9.

Actually there are no obvious laws in Fig. 9. The values of all indices neither increase nor decrease linearly with the increase of threshold. That is because when the thresholds get bigger, more and more high-frequency coefficients are processed using Eq. (29) instead of Eq. (30); maybe some coefficients are canceled out for the opposite signs, and there is more information redundancy in the fused image. So it is difficult to determine how the indices change with the increase of threshold. Hence we set threshold equal to 0.9 in the first and second fusion experiments empirically.

D. Other Fusion Examples

We have tested our algorithm on many datasets; some results are shown here. Because of space limitations, only the results of curvelet with EOG are shown.

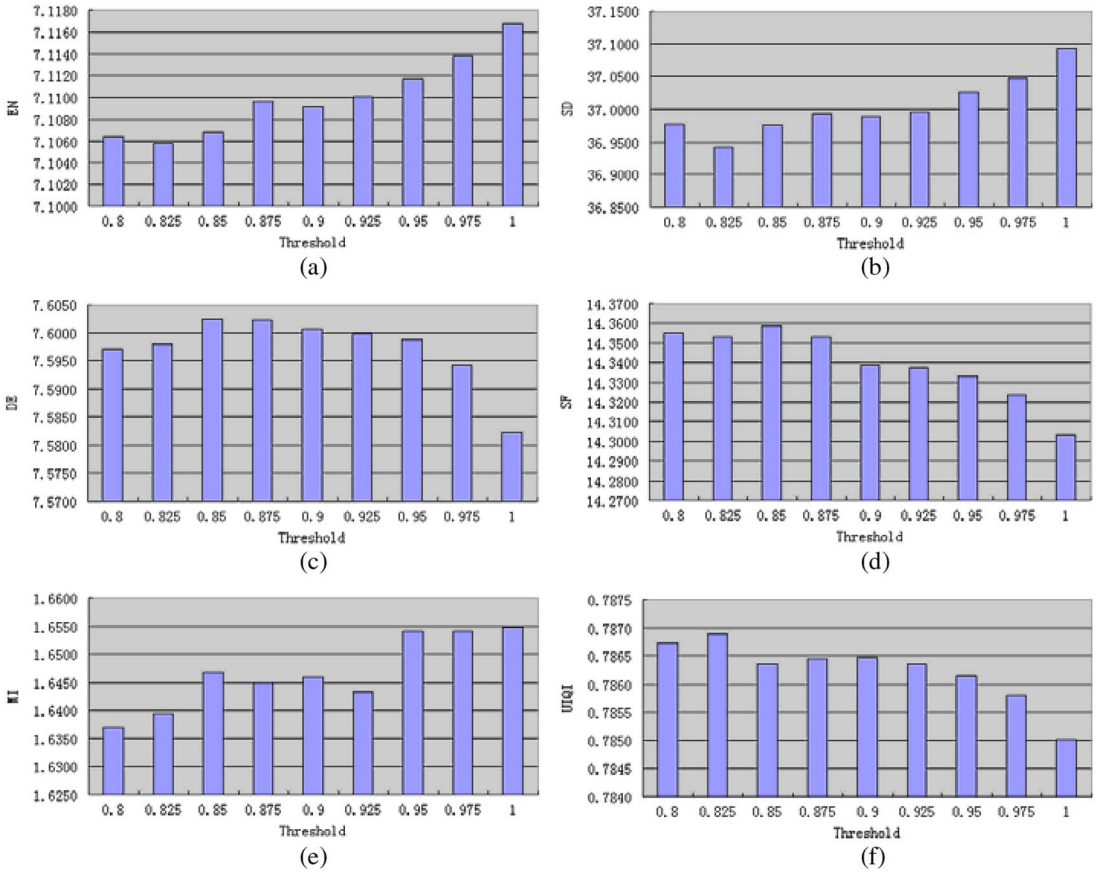

Fig. 9. (Color online) Influence of threshold of FOCC.

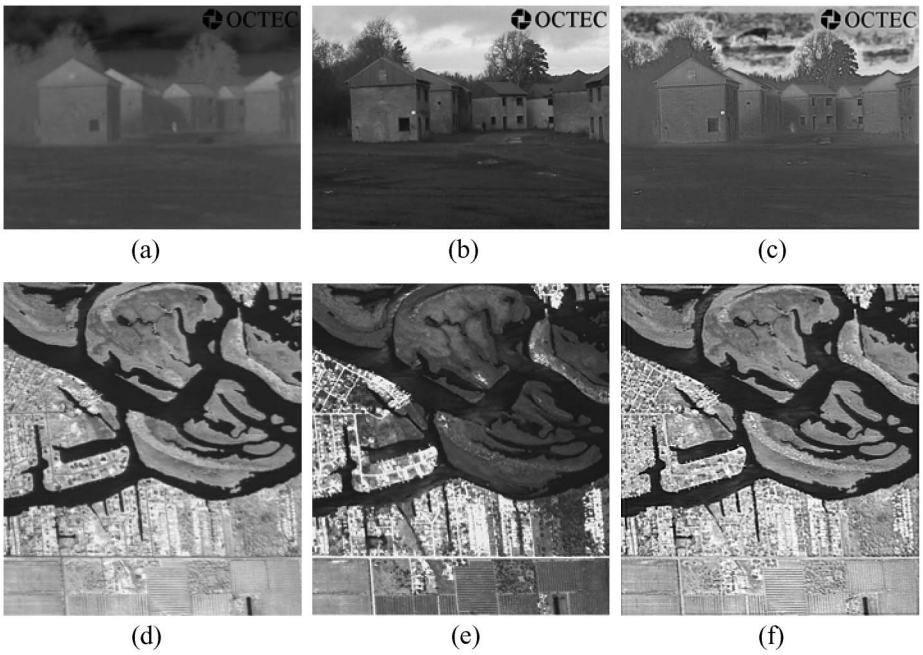

Fig. 10. Results of other datasets: (a) infrared image, (b) visible image, (c) fusion image, (d) infrared remote sensing image, (e) visible remote sensing image, and (f) fusion image.

Figures 10(d) and 10(e) are the examples of remote sensing images. The following results indicate that the image fusion is an effective way to integrate the advantages of infrared and visible images.

6. Conclusion

This paper introduced the focus measure operators into the curvelet domain, and presented a new infrared and visible image fusion method. The local variance weighted strategy is applied to fuse the low-frequency coefficients. Not only are the particular low-frequency coefficients of the visible image injected into the fused image, but also all the low-frequency information is maintained. The high-frequency coefficients are obtained by using FOCC match rules to maintain more detail information of the visible image. The experimental results indicate that the objective evaluation results of proposed methods are close to each other and are superior to the traditional methods in visual effect and performance evaluation.

This work is jointly supported by the National Basic Research Program of China (973 Program, No. 2010CB731800), 863 project (No. 2009AA121404), the National Natural Science Foundation program (No. 61172174, No. 40801165, and No. 10978003), and the Fundamental Research Funds for the Central Universities (No. 111056 and No. 201121302020008).

References

1. F. F. Zhou, W. D. Chen, and L. F. Li, "Fusion of IR and visible images using region growing," J. Appl. Opt. **28**, 737–741 (2007) (in Chinese).
2. A. Apatean, C. Rusu, A. Rogozan, and A. Bensrhair, "Visible-infrared fusion in the frame of an obstacle recognition system," in *IEEE International Conference on Automation, Quality and Testing, Robotics* (IEEE, 2010), pp. 1–6.
3. X. H. Yang, H. Y. Jin, and L. C. Jiao, "Adaptive image fusion algorithm for infrared and visible light images based on DT-CWT," J. Infrared Millim. Waves **26**, 419–424 (2007).
4. H. M. Wang, K. Zhang, and Y. J. Li, "Image fusion algorithm based on wavelet transform," Infrared Laser Eng. **34**, 328–332 (2005) (in Chinese).
5. S. Firooz, "Comparative image fusion analysis," in *IEEE Proceedings of the Conference on Computer Vision and Pattern Recognition* (IEEE, 2005), pp. 1–8.
6. G. Pajares and J. M. de la Cruz, "A wavelet-based image fusion tutorial," Pattern Recogn. **37**, 1855–1872 (2004).
7. R. Minhas, A. A. Mohammed, and Q. M. J. Wu, "Shape from focus using fast discrete curvelet transform," Pattern Recogn. **44**, 839–853 (2011).
8. S. Li and B. Yang, "Multifocus image fusion by combining curvelet and wavelet transform," Pattern Recogn. Lett. **29**, 1295–1301 (2008).
9. X. B. Qu, J. W. Yan, and G. D. Yang, "Multifocus image fusion method of sharp frequency localized contourlet transform domain based on sum-modified-Laplacian," Opt. Precision Eng. **17**, 1203–1212 (2009).
10. E. J. Candès and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with C2 singularities," Commun. Pure Appl. Math. **57**, 219–266 (2004).
11. E. J. Candès, L. Demanet, and D. L. Donoho, "Fast discrete curvelet transforms," Multiscale Model. Simul. **5**, 861–899 (2006).
12. W. Huang and Z. L. Jing, "Evaluation of focus measures in multi-focus image fusion," Pattern Recogn. Lett. **28**, 493–500 (2007).
13. R. Minhas, A. A. Mohammed, and Q. M. J. Wu, "An efficient algorithm for focus measure computation in constant time," IEEE Trans. Circuits Syst. Video Technol. **22**: 152–156 (2012).
14. R. Minhas, A. A. Mohammed, Q. M. J. Wu, and M. A. Sid-Ahmed, "3D shape from focus and depth map computation using steerable filters," Lect. Notes Comput. Sci. **5627**, 573–583 (2009).
15. M. Muhammad and T. S. Choi, "Sampling for shape from focus in optical microscopy," IEEE Trans. Pattern Anal. Machine Intell. **99**, 1–12 (2011).
16. V. Aslantas and R. Kurban, "A comparison of criterion functions for fusion of multi-focus noisy images," Opt. Commun. **282**, 3231–3242 (2009).

17. S. K. Nayar and Y. Nakagawa, "Shape from focus," IEEE Trans. Pattern Anal. Machine Intell. **16**, 824–831 (1994).
18. Z. Wang and A. C. Bovik, "A universal image quality index," IEEE Signal Process Lett. **9**, 81–84 (2002).
19. G. X. Liu, S. G. Zhao, and W. H. Yang, "Multi-sensor image fusion scheme based on gradient pyramid decomposition," J. Optoelectron. Laser **12**, 293–296 (2001) (in Chinese).
20. J. Liu and Z. F. Shao, "Feature-based remote sensing image fusion quality metrics using structure similarity," Acta Photon. Sin. **40**, 126–131 (2011).
21. M. Welling, "Robust higher order statistics," in *Proceedings of 10th International Workshop on Artificial Intelligence and Statistics* (2005), pp. 405–412.