Journal of Computers Vol. 28, No. 2, 2017, pp. 57-71
doi:10.3966/199115592017042802005

Infrared and Visible Images Fusion Method Based on Discrete Wavelet Transform

Lingchao Zhan[1,2], Yi Zhuang[1], and Longda Huang[3]

[1] College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing210016, China
496043955@qq.com,zhuangyi@nuaa.edu.cn

[2] Nanhang Jincheng College, Nanjing211156, China
496043955@qq.com

[3] China Electric Power Research Institute, Nanjing210000, China
6438288@qq.com

Received 07 December 2015; Revised 10 March 2016; Accepted 20 April 2016

Abstract. For infrared images, the performance of texture details in the target scene is not good, and for visible images, the performance is constrained by illumination and environment. In this paper, an image fusion method based on Discrete Wavelet Transform (DWT) is proposed. First, the infrared and visible images are decomposed by DWT, the low frequency and high frequency components will be obtained. For the low frequency component, the fusion method is based on the regional energy. For the high frequency components, the fusion method is based on the weighted sum of the difference between the neighboring coefficients. Finally, the fusion image is obtained by the inverse DWT. The results prove that, compared with the similar research methods, the performance of the proposed method is better.

Keywords: discrete wavelet transform, image fusion, infrared image, visible image

1 Introduction

Image fusion is the method to fuse two or more images into a single image, this fused image should include more information. In recent years, image fusion technology is applied in many fields, such as remote sensing, medical image processing, robot vision and digital camera application.

Infrared and visible image fusion method is one of the image fusion methods. The infrared image is the result of the infrared thermal imaging sensors detecting the outside temperature difference. It can work all day and not easily affected by the environment, but the performance of the image texture details is not good. The visible image is suitable for human visual characteristics, but it is constrained by lighting and environmental conditions. For these characteristics of the two kinds of images, we fuse them together. Not only have the promise of the quality, but also make the image clearer. It can improve the image resolution and enhance the details. It is convenient for human identification.

Image fusion can be divided into the pixel level fusion, feature level fusion and decision level fusion. The pixel level fusion is the fusion of the source image data. It directly fuses the multi source images of the same scene together. The feature level image fusion has two steps. The pretreatment and feature extraction of the source images is the first step, the second step is the fusion of the various features. The decision level fusion is a kind of high level data fusion. First, for the multiple source images, each feature of the same target is classified and judged. Then the independent decisions of each source images will be fused together. Finally, the joint decision will be obtained. Our method is part of the pixel level fusion. At present, the commonly used methods of the pixel level fusion include PCA [1], the Pyramid transform

[2-3], wavelet transform [4-16]. PCA is a method of selecting optimal pixel weight. The disadvantage is it cannot prominent spectral characteristics, so it is not suitable for the weak correlation of image fusion. The Pyramid decomposition aims to decompose images into different spatial frequency. Using the characteristics of the decomposition of the tower structure has different decomposition layers in different spatial resolution, it adopts different fusion operators. Finally, the characteristics and details which are from different images will be fused together. But it leads to bad fusion result because of the correlation between the layers.

The wavelet transform has good performance in preserving image information. Different results can be obtained through different fusion rules. In this paper, a fusion method for the infrared and visible images is proposed based on discrete wavelet transform. The two images to be fused are decomposed by DWT, the low frequency and high frequency components will be obtained. For the low frequency component, the fusion method is based on the regional energy. For the high frequency components, the fusion method is based on the weighted sum of the difference between the neighboring coefficients. Finally, the fusion image is obtained by the inverse DWT. The fusion results of the proposed algorithm are compared with the results of several other algorithms. We use five objective evaluation standards, they are mean, information entropy, standard deviation, average gradient and mutual information. The results show that our method has a good effect, our fusion results are clearer.

2 Related Research

2.1 Related Research Based on Wavelet Transform

Wavelet transform was first proposed by the French scientist Morlet and Grossman in the early 1980s. With the rise of wavelet theory, especially after scientist Mallat proposed the fast wavelet transform algorithm [5], wavelet transform has been widely applied in the field of image processing. Now it has been a typical image fusion tool. Compared with the pyramid transform, decomposition coefficients between different resolutions are not relevant.

In the early 1990s, Li et al. [10], Chipman and Orr [11] first proposed the concept of image fusion based on wavelet analysis. With the continuous development of wavelet theory and the requirements of improving the image fusion, image fusion is the methods to fuse the wavelet coefficients based on different fusion rules [4, 7, 12]. Chang et al. and Yang et al. are developed by taking into account the characteristics of human visual system [6, 9]. Wang and Zheng [8] proposed a fusion method based on YUV color space and wavelet transform. Li et al. [13] proposed a fusion method based on wavelet transform and edge keeping method. Kingsbury proposed dual tree complex wavelet transform has better direction selectivity and approximate translation invariance [14]. Tian and Chen proposed a statistical sharpness measure. Singh and Khare use Daubechies complex wavelet transform (DCxWT) for image fusion. In this paper, a new fusion method based on discrete wavelet transform is proposed for the fusion of visible and infrared images [16].

2.2 Fusion Method Based on Wavelet Transform

For the inspiration of the tower algorithm for image decomposition and reconstruction, Mallat [5] proposed wavelet decomposition and reconstruction algorithm based on multi-resolution theory, that is the Mallat fast algorithm. The status of the Mallat fast algorithm in wavelet transform is like the fast Fourier transform (FFT) in Fourier transform, it is the basis of wavelet analysis and synthesis.

Suppose the image is $f(x,y)$, the discrete approximation is $A_j^d f$, the discrete detail is $D_j f$. The wavelet basis function is $\psi_{i,j}(x)$, its scaling function is $\phi(x)$. We have the formulae shown in equation (1)-(4) [10].

$$A_j^d f(n,m) = < f(x,y), \phi_{j,n}(x)\phi_{j,m}(y) > \quad \textbf{(1)}$$

$$D_j^1 f(n,m) = < f(x,y), \phi_{j,n}(x)\psi_{j,m}(y) > \quad \textbf{(2)}$$

$$D_j^2 f(n,m) =< f(x,y), \psi_{j,n}(x)\phi_{j,m}(y) > \quad (3)$$

$$D_j^3 f(n,m) =< f(x,y), \psi_{j,n}(x)\psi_{j,m}(y) > \quad (4)$$

The decomposition formulae are shown in equation (5)-(8) [10].

$$A_j^d f(n,m) = \sum_k \sum_l h_{2n-k} h_{2m-1} A_{j+1}^d f(k,l) \quad (5)$$

$$D_j^1 f(n,m) = \sum_k \sum_l h_{2n-k} g_{2m-1} A_{j+1}^d f(k,l) \quad (6)$$

$$D_j^2 f(n,m) = \sum_k \sum_l g_{2n-k} h_{2m-1} A_{j+1}^d f(k,l) \quad (7)$$

$$D_j^3 f(n,m) = \sum_k \sum_l g_{2n-k} g_{2m-1} A_{j+1}^d f(k,l) \quad (8)$$

The image is decomposed in the following way: in each layer of the transformation, the image is decomposed into four smaller images, that is 1/4 size of the original image.

Each of the four images are generated by 2 times the interval sampling in both of the X and Y directions after the inner product of the original image and a wavelet basis.

The corresponding reconstruction formula is shown in (9).

$$\begin{aligned} A_{j+1}^d f(n,m) = &\sum_k \sum_l h_{n-2k} h_{m-2l} A_j^d f(k,l) + \sum_k \sum_l h_{n-2k} g_{m-2l} D_j^1 f(k,l) \\ &+ \sum_k \sum_l g_{n-2k} h_{m-2l} D_j^2 f(k,l) + \sum_k \sum_l g_{n-2k} g_{m-2l} D_j^3 f(k,l) \end{aligned} \quad (9)$$

When the image is decomposed and reconstructed, the filter banks are h and g, the impulse response functions are h_k and g_k. They meet the following conditions:

1. Orthogonality:

$$\sum_{k \in z} h_k h_{k+2n} = \delta(0,2n) \quad (10)$$

2. Normalization:

$$\sum_{k \in z} h_k = \sqrt{2}, \sum_{k \in z} g_k = 0 \quad (11)$$

3.

$$g_k = (-1)^k h_{1-k} \quad (12)$$

4.

$$h_k = h_{-k} \quad g_k = g_{-k} \quad (13)$$

When the signal of sample is limited, the formulae are shown in (14).

$$h_k = h_{1-k} \quad g_k = g_{1-k} \quad (14)$$

For the image $f(n,m)$, n is the line number, corresponding to the sampling of x direction, m is the column number, corresponding to the sampling of y direction. The following conclusions can be obtained:

1. $A_j^d f$ is the approximate image. That is the original image been filtered out the high frequency in both the row and column directions.

2. $D_j^1 f$ is the high frequency component in the vertical direction. That is the original image been filtered out the high frequency in the row direction and the low frequency in the column direction.

3. $D_j^2 f$ is the high frequency component in the horizontal direction. That is the original image been filtered out the low frequency in the row direction and the high frequency in the column direction.

4. $D_j^3 f$ is the high frequency component in the diagonal direction. That is the original image been filtered out the low frequency in both the row and column directions.

3 IVFDWT Method

3.1 The Motivation of the IVFDWT Rules

The infrared images can reflect the existing characteristics of the target, but its visual effect is not good and cannot distinguish image details. The visible images can have a high resolution, but in the case of sheltered or poor lighting, the visible image will have low contrast.

For the characteristics of the infrared and visible images and the wavelet transform really has good performance in image fusion [4-16], a fusion method named IVFDWT is proposed in this paper. This method is Infrared and Visible Images Fusion Method Based on Discrete Wavelet Transform.

3.2 The Design of the IVFDWT Rules

The more levels of the wavelet transform, the more details in the fusion result. But with the increase of decomposition levels, the image structural information loss will become larger, this information will not be recovered in the inverse wavelet transform. So the wavelet decomposition level cannot be too large, generally between 2 to 5. It has proved by experiments and we consider the running time, this paper choose 2 as the decomposition level.

Definition 1. The two images to be fused are image A and image B. They are decomposed using discrete wavelet transform. After the decomposition, the low frequency component of image A is $L_A(x,y)$ and the low frequency component of image B is $L_B(x,y)$. The high frequency component are represented by $H_A^{h1}(x,y)$, $H_A^{v1}(x,y)$, $H_A^{d1}(x,y)$, $H_A^{h2}(x,y)$, $H_A^{v2}(x,y)$, $H_A^{d2}(x,y)$, $H_B^{h1}(x,y)$, $H_B^{v1}(x,y)$, $H_B^{d1}(x,y)$, $H_B^{h2}(x,y)$, $H_B^{v2}(x,y)$, $H_B^{d2}(x,y)$, they are the horizontal, vertical and diagonal high frequency components of the first and the second level, respectively.

The DWT decomposition diagrams of image A and B are shown in Fig. 1.

(a) decomposition diagram of image A (b) decomposition diagram of image B

Fig. 1. The DWT decomposition diagrams of image A and B

The low and high frequency components of the image A and B will be fused using different fusion rules. Our rules are as follows:

(1) For the low frequency coefficient, the result is determined by the regional energy. The regional energy of the low frequency coefficient will be calculated, then the fusion result of the low frequency coefficient will be obtained according to the regional energy and the corresponding correlation value. The specific description is in Section 3.3.

(2) For the high frequency coefficient. The difference of neighboring coefficients will first be calculated, then the weighted sum of the difference will be obtained. If the weighted sum of image A is greater

than that of image B, the fusion result is the high frequency coefficient of image A, otherwise, select the high frequency coefficient of image B.

The fused image will be obtained by the inverse DWT.

3.3 The Description of IVFDWT Method

The two images to be fused are image A and B, the diagram of IVFDWT method is shown in Fig. 2.

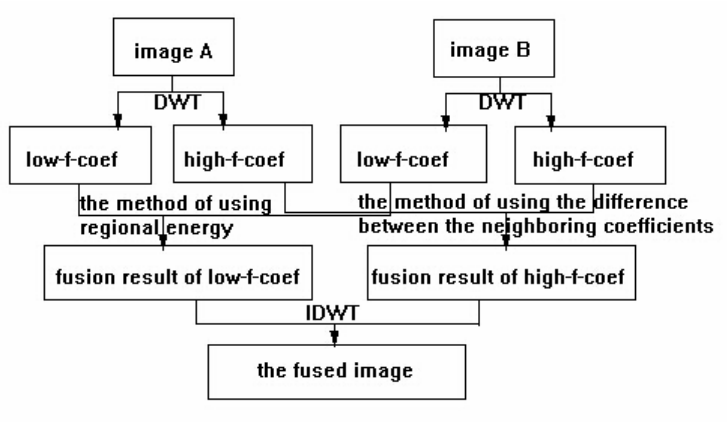


Fig. 2. The diagram of IVFDWT method

The two images to be fused are first decomposed using Discrete Wavelet Transform (DWT). The low frequency coefficients and the high frequency coefficients are fused using different fusion rules.

The IVFDWT method is described as follows:

The fusion method of the low frequency coefficients. The low frequency components of the two images are divided into 3×3 regions. (we have tried the dimensions of 3×3 and 5×5, experiments show that the 3×3 is better. The following are all divided into 3×3 regions.) Then we calculate each regional energy. Suppose E_A and E_B represent the regional energy of the image A and image B. The energy formulae are shown in equation (15)-(16) [7].

$$E_A(x,y)=\sum_{i,j=-\frac{N}{2}}^{\frac{N}{2}}\omega(i+2,j+2)(L_A(x+i,y+j))^2\log((L_A(x+i,y+j))^2) \tag{15}$$

$$E_B(x,y)=\sum_{i,j=-\frac{N}{2}}^{\frac{N}{2}}\omega(i+2,j+2)(L_B(x+i,y+j))^2\log((L_B(x+i,y+j))^2) \tag{16}$$

where $N=3$, $\omega(x,y)$ is the weight matrix, the formula is shown in (17) [17].

$$\omega=\frac{1}{16}\begin{bmatrix}1&2&1\\2&4&2\\1&2&1\end{bmatrix} \tag{17}$$

According to the energy formulae, we can get an initial fusion decision map. The formula is shown in equation (18).

$$i_L\_map(x,y)=\begin{cases}1 & E_A(x,y)\geq E_B(x,y)\\0 & E_A(x,y)<E_B(x,y)\end{cases} \tag{18}$$

The fusion rule is that if the regional energy of image A is greater than that of image B, the corresponding value of the decision map is 1, otherwise 0.

The initial fusion decision map is changed by the consistency verification by using majority filtering operation. For the initial fusion decision map, in each 3×3 window, if the center value is 0, most of the other values are also 0, then the center value keeps the same. If the center value is 0, most of the other

values are 1, then the center value will be 1.

The correlation of the low frequency coefficients is calculated according to the energy value. The correlation is the relationship between the corresponding regions of the low frequency coefficients. The formula for calculating the correlation is shown in equation (19) [4].

$$M(x,y)=\frac{2\times\sum_{i=-\frac{N}{2}}^{\frac{N}{2}}\sum_{j=-\frac{N}{2}}^{\frac{N}{2}}L_A(x+i,y+j)\times L_B(x+i,y+j)}{E_A(x,y)+E_B(x,y)} \tag{19}$$

where $N=3$, E_A and E_B are the corresponding regional energies of the low frequency coefficients. The value of M in (19) is in the range of 0~1, we choose 0.65 as the threshold. (we have tried 0.6,0.65 and 0.7 as the threshold, experiments show that 0.65 is better.) If $M<0.65$, the decision map keeps the same. If $M\geq 0.65$, the value of the corresponding position in the fusion map will be 2. Then we get the final fusion decision map $f_L\_map(x,y)$.

We use the final fusion decision map to get the fusion result of the low frequency coefficients. The formula is shown in (20).

$$L(x,y)=\begin{cases}\dfrac{E_A(x,y)}{E_A(x,y)+E_B(x,y)}L_A(x,y)+\dfrac{E_B(x,y)}{E_A(x,y)+E_B(x,y)}L_B(x,y) & \\ & f_L\_map(x,y)=2\\ L_A(x,y) & f_L\_map(x,y)=1\\ L_B(x,y) & f_L\_map(x,y)=0\end{cases} \tag{20}$$

If the value in the final fusion decision map is 2, the fusion result is the weighted sum of the two images' low frequency coefficients. If the value in the map is 1, the fusion result is $L_A(x,y)$. If the value in the map is 0, the fusion result is $L_B(x,y)$.

The fusion method of the high frequency coefficients. The high frequency coefficients reflect the edge and detail information of an image, the neighboring coefficients are fully considered here. The fusion rule is based on the weighted sum of the difference between the neighboring coefficients. First, the difference between the neighboring coefficients is calculated, the formula is shown in equation (21) [17].

$$\begin{aligned}M(x,y)&=\left|2H(x,y)-H(x-1,y)-H(x+1,y)\right|\\&+|\,2H(x,y)-H(x,y-1)-H(x,y+1)\,|+\\&|\,4H(x,y)-H(x-1,y-1)-H(x-1,y+1)-H(x+1,y-1)-H(x+1,y+1)\,|\end{aligned} \tag{21}$$

the formula of the weighted sum is shown in equation (22).

$$WM(x,y)=\sum_{i,j=-\frac{N}{2}}^{\frac{N}{2}}\omega(i+2,j+2)M(x+i,y+j) \tag{22}$$

Where N=3, $\omega(x,y)$ is shown in (17).

According to equation (21)-(22), an initial fusion decision map will be obtained. The formulae are shown in equation (23)-(28).

$$i_H^{h1}\_map(x,y)=\begin{cases}1 & WM_A^{h1}(x,y)\geq WM_B^{h1}(x,y)\\0 & WM_A^{h1}(x,y)<WM_B^{h1}(x,y)\end{cases} \tag{23}$$

$$i_H^{v1}\_map(x,y)=\begin{cases}1 & WM_A^{v1}(x,y)\geq WM_B^{v1}(x,y)\\0 & WM_A^{v1}(x,y)<WM_B^{v1}(x,y)\end{cases} \tag{24}$$

$$i_H^{d1}\_map(x,y)=\begin{cases}1 & WM_A^{d1}(x,y)\geq WM_B^{d1}(x,y)\\ 0 & WM_A^{d1}(x,y)<WM_B^{d1}(x,y)\end{cases} \tag{25}$$

$$i_H^{h2}\_map(x,y)=\begin{cases}1 & WM_A^{h2}(x,y)\geq WM_B^{h2}(x,y)\\ 0 & WM_A^{h2}(x,y)<WM_B^{h2}(x,y)\end{cases} \tag{26}$$

$$i_H^{v2}\_map(x,y)=\begin{cases}1 & WM_A^{v2}(x,y)\geq WM_B^{v2}(x,y)\\ 0 & WM_A^{v2}(x,y)<WM_B^{v2}(x,y)\end{cases} \tag{27}$$

$$i_H^{d2}\_map(x,y)=\begin{cases}1 & WM_A^{d2}(x,y)\geq WM_B^{d2}(x,y)\\ 0 & WM_A^{d2}(x,y)<WM_B^{d2}(x,y)\end{cases} \tag{28}$$

The fusion rule is that if the weighted sum of image A is greater than that of image B, the corresponding value of the decision map is 1, otherwise 0.

The final fusion decision maps are also obtained by the consistency verification by using majority filtering operation. For the initial fusion decision maps, in each 3×3 window, if the center value is 0, most of the other values are also 0, then the center value keeps the same. If the center value is 0, most of the other values are 1, then the center value will be 1.

We use the final fusion decision map to get the fusion result of the high frequency coefficients. The formulae are shown in equation (29)-(34).

$$H^{h1}(x,y)=\begin{cases}H_A^{h1}(x,y) & f_H^{h1}\_map(x,y)=1\\ H_B^{h1}(x,y) & f_H^{h1}\_map(x,y)=0\end{cases} \tag{29}$$

$$H^{v1}(x,y)=\begin{cases}H_A^{v1}(x,y) & f_H^{v1}\_map(x,y)=1\\ H_B^{v1}(x,y) & f_H^{v1}\_map(x,y)=0\end{cases} \tag{30}$$

$$H^{d1}(x,y)=\begin{cases}H_A^{d1}(x,y) & f_H^{d1}\_map(x,y)=1\\ H_B^{d1}(x,y) & f_H^{d1}\_map(x,y)=0\end{cases} \tag{31}$$

$$H^{h2}(x,y)=\begin{cases}H_A^{h2}(x,y) & f_H^{h2}\_map(x,y)=1\\ H_B^{h2}(x,y) & f_H^{h2}\_map(x,y)=0\end{cases} \tag{32}$$

$$H^{v2}(x,y)=\begin{cases}H_A^{v2}(x,y) & f_H^{v2}\_map(x,y)=1\\ H_B^{v2}(x,y) & f_H^{v2}\_map(x,y)=0\end{cases} \tag{33}$$

$$H^{d2}(x,y)=\begin{cases}H_A^{d2}(x,y) & f_H^{d2}\_map(x,y)=1\\ H_B^{d2}(x,y) & f_H^{d2}\_map(x,y)=0\end{cases} \tag{34}$$

If the value in the map is 1, the fusion result is from A. If the value in the map is 0, the fusion result is from B. The inverse DWT will be taken to the resulting low frequency and high frequency coefficients, the final fusion result will be obtained.

4 Experiments and Analysis

4.1 Evaluation Criteria

In order to evaluate the performance of the proposed method, we adopt mean, information entropy (IE), standard deviation (SD), average gradient (AG) and mutual information (MI) as objective evaluation criteria to evaluate the results of our fusion method.

The mean is the mean value of the image pixel values. It reflects the average brightness of an image. The calculation formula is shown in equation (35) [4].

$$\bar{I} = \frac{1}{m \times n} \sum_{x=1}^{m} \sum_{y=1}^{n} I(x, y) \tag{35}$$

$I(x, y)$ is the pixel value of an image, $m \times n$ is the size of an image.

Information entropy is an important indicator of the evaluation of images. The calculation formula is shown in equation (36) [21].

$$H = -\sum_{i=0}^{L-1} P_i \log_2 P_i \tag{36}$$

P_i is the distribution probability of the gray value i, L is the total number of the gray level.

Standard deviation reflects the image texture information. The larger the standard deviation, the more dispersed distribution of the gray levels of an image, the sharper textures. The calculation formula is shown in equation (37) [21].

$$SD = \sqrt{\frac{1}{m \times n} \sum_{x=1}^{m} \sum_{y=1}^{n} (I(x, y) - \bar{I})^2} \tag{37}$$

Average gradient reflects the tiny details of variance, texture variation and image resolution. The greater the average gradient, the better the image resolution. The calculation formula is shown in equation (38) [21].

$$AG = \frac{1}{mn} \sum_{x=1}^{m} \sum_{y=1}^{n} \sqrt{\frac{[I(x+1, y) - I(x, y)]^2 + [I(x, y+1) - I(x, y)]^2}{2}} \tag{38}$$

$I(x, y)$ is the pixel value of an image, $m \times n$ is the size of an image.

Mutual information is a measure of the shared information between two images. The greater the value of mutual information, the better the effect of the fused image. Suppose the two images to be fused are A and B, the fused image is F. The calculation formulae of the mutual information that F contains about A and B are shown in equation (39)-(40) [22].

$$I_{FA}(f, a) = \sum_{f,a} P_{FA}(f, a) \log \frac{P_{FA}(f, a)}{P_F(f) P_A(a)} \tag{39}$$

$$I_{FB}(f, b) = \sum_{f,b} P_{FB}(f, b) \log \frac{P_{FB}(f, b)}{P_F(f) P_B(b)} \tag{40}$$

The mutual information of the image fusion performance measure can be calculated in equation (41) [22].

$$MI = I_{FA}(f, a) + I_{FB}(f, b) \tag{41}$$

4.2 Fusion Results

In this section, two different infrared and visible image datasets are provided to demonstrate the performance of the proposed algorithm. There are three groups of contrast experiments: (1) pixel averaging method (P_avg), (2) discrete wavelet transform fusion method (DWT) [8], (3) NSCT fusion method [17]. The method is experimented in MATLAB environment.

The infrared and visible images in the experiment have been matched exactly, they were shot in the same scene and at the same time. Fig. 3 shows one pair of images from the first dataset, (a) is an infrared image, (b) is a visible image. Fig. 4 shows four pairs of images from the second dataset, (a1) (a2) (a3) (a4) are infrared images, (b1) (b2) (b3) (b4) are visible images.

(a) (b)

Fig. 3. Infrared and visible images: (a) infrared image, (b) visible image

(a) (b)

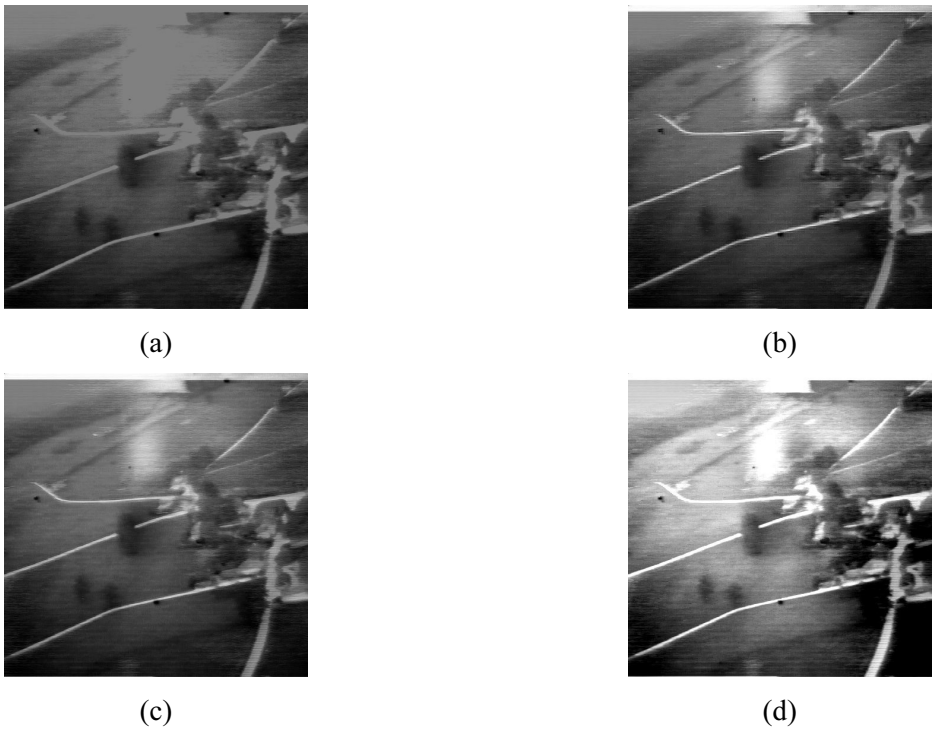
(a2) (b2)

(a3) (b3)

(a4) (b4)

Fig. 4. Infrared and visible images: (a1)-(a4) infrared images, (b1)-(b4) visible images

The results of Fig. 3 are shown in Fig. 5. The three groups of contrast experiments: (1) P_avg is shown in Fig. 5 (a), (2) DWT is shown in Fig. 5 (b), (3) NSCT is shown in Fig. 5 (c). The result of our method is shown in Fig. 5 (d). The p_avg method simply takes the average pixel value of the infrared and visible images, the visual effect is not very good, the effects of the DWT and NSCT methods are better, the visual effect of our method is the best.



(a) (b)

(c) (d)

Fig. 5. Fusion results of figure 3: (a) fused image by P_avg, (b) fused image by DWT, (c) fused image by NSCT, (d) fused image by the proposed method

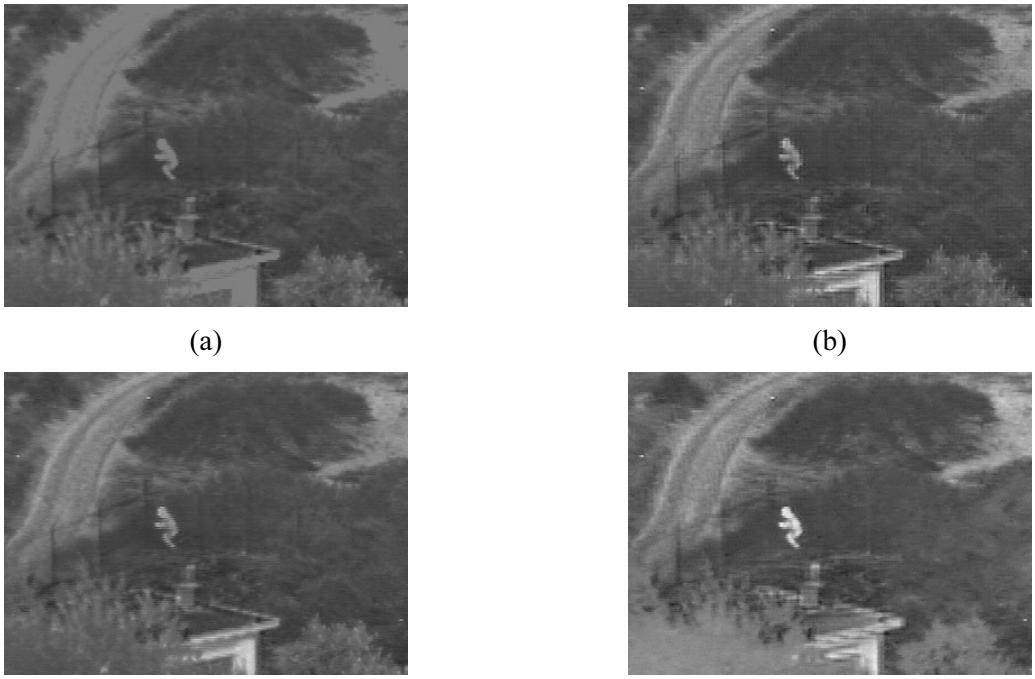
The fusion results of the second dataset are shown in Fig. 6, Fig. 7, Fig. 8 and Fig. 9. The results of the first pair (Fig. 4 (a1) and Fig. 4 (b1)) are shown in Fig. 6. The three groups of contrast experiments: (1) P_avg is shown in Fig. 6 (a), (2) DWT is shown in Fig. 6 (b), (3) NSCT is shown in figure 6 (c). The result of our method is shown in Fig. 6 (d).

The results of the second pair (Fig. 4 (a2) and Fig. 4 (b2)) are shown in Fig. 7. The three groups of contrast experiments: (1) P_avg is shown in figure 7(a), (2) DWT is shown in Fig. 7 (b), (3) NSCT is shown in Fig. 7 (c). The result of our method is shown in Fig. 7 (d).

The results of the third pair (Fig. 4 (a3) and Fig. 4 (b3)) are shown in Fig. 8. The three groups of contrast experiments: (1) P_avg is shown in Fig. 8 (a), (2) DWT is shown in Fig. 8 (b), (3) NSCT is shown in Fig. 8 (c). The result of our method is shown in Fig. 8 (d).

The results of the fourth pair (Fig. 4 (a4) and Fig. 4 (b4)) are shown in Fig. 9. The three groups of contrast experiments: (1) P_avg is shown in figure 9 (a), (2) DWT is shown in Fig. 9 (b), (3) NSCT is shown in Fig. 9 (c). The result of our method is shown in Fig. 9 (d).

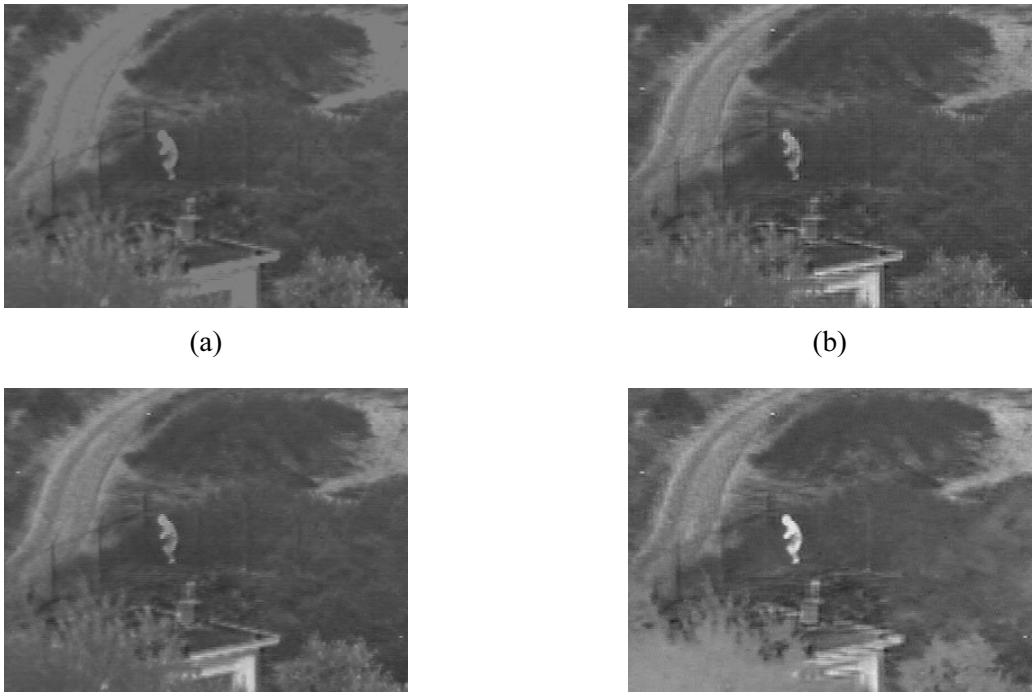
We can see from the fusion results of the above four pairs of images, the visual effect of our method is the best.



(a) (b)

(c) (d)

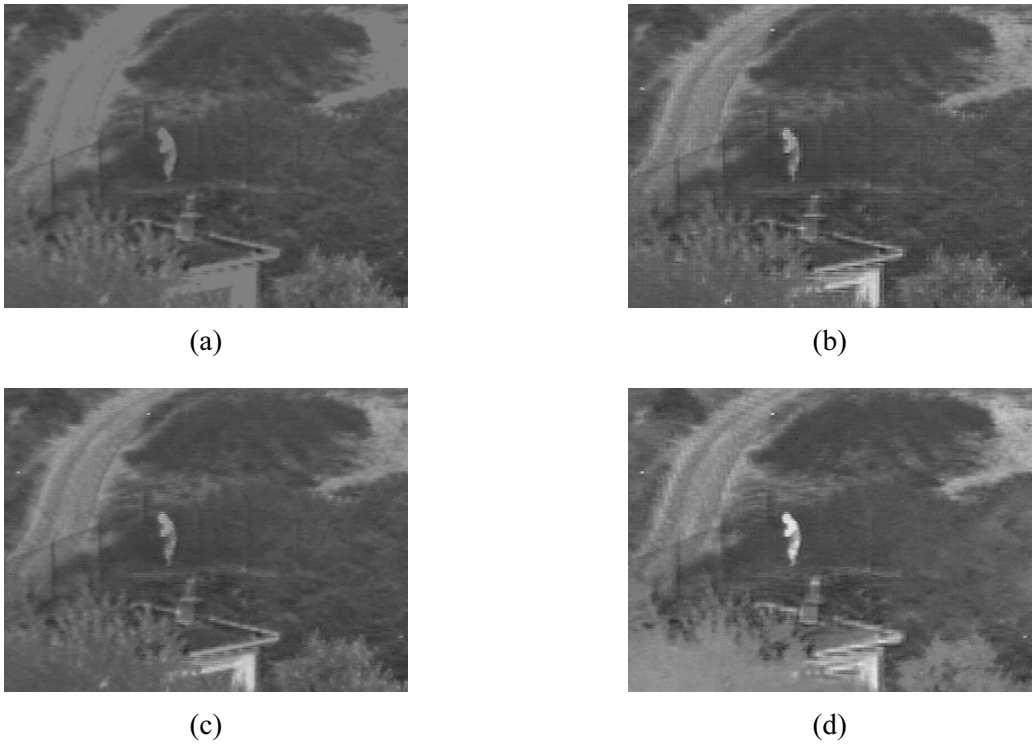
Fig. 6. Fusion results (the first pair): (a) fused image by P_avg, (b) fused image by DWT, (c) fused image by NSCT, (d) fused image by the proposed method



(a) (b)

(c) (d)

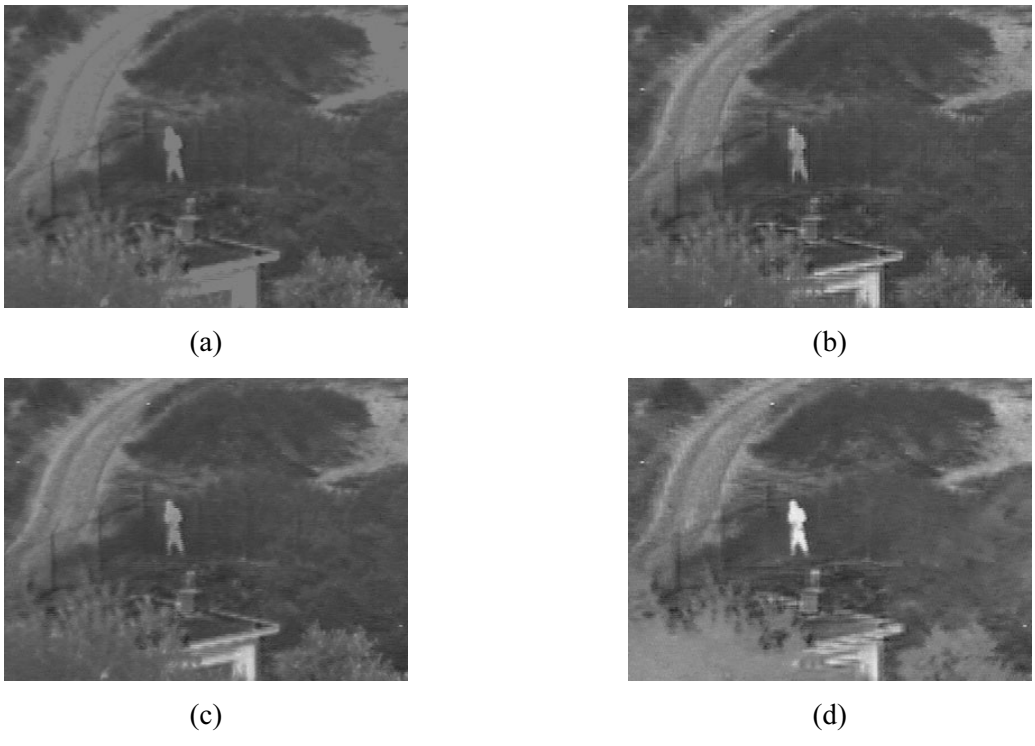
Fig. 7. Fusion results (the second pair): (a) fused image by P_avg, (b) fused image by DWT_avg_max, (c) fused image by NSCT_avg_max, (d) fused image by the proposed method



(a) (b)

(c) (d)

Fig. 8. Fusion results (the third pair): (a) fused image by P_avg, (b) fused image by DWT_avg_max, (c) fused image by NSCT_avg_max, (d) fused image by the proposed method



(a) (b)

(c) (d)

Fig. 9. Fusion results (the fourth pair): (a) fused image by P_avg, (b) fused image by DWT_avg_max, (c) fused image by NSCT_avg_max, (d) fused image by the proposed method

4.3 The Analysis

The fusion results are evaluated by the evaluation criteria mentioned in 4.1. The evaluation of the fusion results of the first dataset is shown in Table 1, the evaluation of the fusion results of the second dataset is shown in Table 2, Table 2, Table 3, Table 4 and Table 5.

Table 1. The evaluation of the fusion results of the first dataset

Evaluation Criteria	P_avg	DWT	NSCT	proposed method
MEAN	78.9623	84.3814	84.3786	**116.2095**
IE	6.4524	7.4837	7.4678	**7.9436**
SD	36.4378	47.9245	47.6339	**58.8363**
AG	4.9085	9.1209	7.4678	**9.1559**
MI	4.0651	5.0892	5.3171	**5.4593**

Table 2. The evaluation of the fusion results of the second dataset (the first pair)

Evaluation Criteria	P_avg	DWT	NSCT	proposed method
MEAN	89.3668	90.5426	90.5477	**102.1073**
IE	5.7444	6.3229	6.2857	**6.8216**
SD	19.6530	23.7689	23.3830	**28.1267**
AG	4.3066	7.4081	6.4411	**7.6720**
MI	2.1841	3.9072	4.0390	**4.5371**

Table 3. The evaluation of the fusion results of the second dataset (the second pair)

Evaluation Criteria	P_avg	DWT	NSCT	proposed method
MEAN	90.4626	91.5878	91.5929	**103.7606**
IE	5.7076	6.2769	6.2390	**6.5485**
SD	19.0572	23.0303	22.6538	**26.5359**
AG	4.1075	**7.3024**	6.1195	7.2208
MI	2.1261	3.8872	4.0281	**4.4681**

Table 4. The evaluation of the fusion results of the second dataset (the third pair)

Evaluation Criteria	P_avg	DWT	NSCT	proposed method
MEAN	91.1605	92.4325	92.4374	**103.7950**
IE	5.6828	6.2981	6.2631	**6.6902**
SD	19.0475	23.2823	22.9191	**27.5008**
AG	4.2607	7.3469	6.3318	**7.6253**
MI	2.1061	3.9591	4.1402	**4.7771**

Table 5. The evaluation of the fusion results of the second dataset (the fourth pair)

Evaluation Criteria	P_avg	DWT	NSCT	proposed method
MEAN	88.9252	90.0091	90.0132	**102.3968**
IE	5.7600	6.3106	6.2775	**6.6070**
SD	19.6632	23.5516	23.1787	**27.7014**
AG	4.2970	7.2936	6.3487	**7.3570**
MI	2.0551	3.7552	3.9001	**4.5962**

As can be seen from Table 1, compared with the other three fusion methods, the performance of our method is better. As can be seen from Table 2, Table 3, Table 4 and Table 5, compared with the other three fusion methods, the performance of our method is better. The proposed method is optimal except the AG value of the second pair in Table 3.

5 Conclusions

The wavelet transform fusion method can get good fusion result. In this paper, according to the characteristics of infrared and visible images, we have proposed a new fusion method based on discrete wavelet transform. For the low frequency component, the fusion method is based on the regional energy. For the high frequency components, the fusion method is based on the weighted sum of the difference between the neighboring coefficients. Two different infrared and visible image datasets have been provided here. Three groups of similar research methods and five objective evaluation criteria have been applied to evaluate the results of the fusion method.

The results show that no matter from the perspective of the visual effect, or from the objective evaluation of fusion results, compared with the similar research methods, the results of our method are better, clearer. We have achieved more satisfactory results.

Acknowledgements

The authors would like to thank the anonymous reviewers and editors for their invaluable suggestions. The work was jointly supported by the National Natural Science Foundation of China (General Program)(NO.61572253), General Program of University Natural Science of Jiangsu Province (NO. 14KJD510004).

References

[1] M. Kumar, S. Dass, A total variation-based algorithm for pixel-level image fusion, IEEE Transactions on Image Procession 9(18)(2009) 2137-2143.

[2] P.J. Burt, A gradient pyramid basis for pattern selective image fusion, in: Proc. the Society for Information Display Conference, 1992.

[3] D.R. Barron, O.D.J. Thomas, Image fusion though consideration of texture components, IEEE Transactions on Electronics Letters 37(12)(2001) 746-748.

[4] H. Chen, Y. Wang, Study for image fusion based on wavelet transform, Microelectronics & Computer 5(27)(2010) 39-41.

[5] S.G. Mallat, A theory for multi-resolution signal decomposition the wavelet representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 11(7)(1989) 647-693.

[6] H.W. Chang, S.D. Lan, Image fusion based on addition of wavelet coefficients, in: Proc. International Conference on Wavelet Analysis and Pattern Recognition, 2007.

[7] J.J. Lewis, R.J. O'Callaghan, Pixel-and region-based image fusion with complex wavelets, Information Fusion 8(2)(2007) 119-130.

[8] J. Wang, S. Zheng, Visible and Infrared image fusion method based on wavelet transform and YUV, Journal of Northwestern Poly Technical University 3(33)(2013) 208-211.

[9] Y. Yang, D.S. Park, S. Huang, N. Rao, Medical image fusion via an effective wavelet-based approach, EURASIP Journal on Advances in Signal Processing 2010(2010) 44-1~44-13.

[10] H. Li, B.S. Manjunath, S.K. Mitra, Multisensor image fusion using the wavelet transform, Graphical Models and Image Processing 3(27)(1995) 235-244.

[11] L.J. Chipman, T.M. Orr (Eds.), Wavelets and images fusion, in: Proc. IEEE International Conference on Image Processing, 1994.

[12] S. Roy, T. Howlader, S.M. Rahman, Image fusion technique using multivariate statistical model for wavelet coefficients,

Signal, Image and Video Processing 2(7)(2013) 355-365.

[13] F. Li, Y. Zhang, Z. Zhou, D.P. Semanek, S. Wang, L. Wu, An improved image fusion algorithm based on wavelet decomposition, Journal of Convergence Information Technology 10(5)(2010) 15-21.

[14] N. Kingsbury, The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters, in: Proc. 8th IEEE Digital Signal Processing, 1998.

[15] J. Tian, L. Chen, Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure, in: Proc. Signal Process92, 2012.

[16] R. Singh, A. Khare, Fusion of multimodal medical images using Daubechies complex wavelet transform—A multi resolution approach, Information Fusion 19(2014) 49-60.

[17] P. Ganasala, V. Kumar, CT and MR image fusion scheme in nonsubsampled contourlet transform domain, Journal of Digital Imaging, 27(3)(2014) 407-418.

[18] Q. Zhang, B. L. Guo, Multifocus image fusion using the nonsubsampled contourlet transform, Signal Process 89(7)(2009) 1334-1346.

[19] X. Qu, J. Yan, H. Xiao, Z. Zhu, Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain, Acta Automatica Sinica 8(2)(2008) 18-38.

[20] G. Piella, New quality measures for image fusion, in: Proc. International Conference on Information Fusion, 2004.

[21] J. Li, J. Sun, X. Mao, Multi resolution fusion of remote sensing images based on resolution degradation model, Geo-Spatial Information Science 8(1)(2005) 50-56.

[22] G. Bhatnagar, Q.M.J. Wu, Z. Liu, Directive contrast based multimodal medical image fusion in NSCT domain, IEEE Transactions On Multimedia 15(5)(2013) 1520-9210.