

# Multi-sensor image fusion based on contrast and directional features optimization

International Journal of Distributed  
Sensor Networks  
2018, Vol. 14(12)  
© The Author(s) 2018  
DOI: 10.1177/1550147718815841  
journals.sagepub.com/home/dsn  


Haiyan Jin<sup>1,2</sup>, Meng Zhang<sup>1</sup>, Zhaolin Xiao<sup>1</sup> and Yanning Li<sup>1</sup>

## Abstract

Multi-sensor image fusion is always an important and opening problem, which can enhance visual quality and benefit some social security applications. In this article, we use contrast pyramid to decompose visible and infrared images, respectively, and the directional filter banks are applied to obtain multiple directional sub-band image features. Then, we compute the decomposition coefficients of visible and infrared images using a low-pass filter on the decomposed data; and finally, we introduce the whale optimization algorithm to search optimal coefficients to reconstruct the final fusion image. The experiments are conducted on multiple datasets with subjective and objective comparisons, in which the qualitative and quantitative analyses indicate the validity of the proposed method.

## Keywords

Image fusion, multi-sensors, directional features, optimization, whale optimization algorithm

Date received: 17 August 2018; accepted: 19 October 2018

Handling Editor: Shancang Li

## Introduction

Traditional imaging device can only capture a limited spectral emission or radiation of lights of the target scene, in which exist unpredictable objects or materials. Compound imaging and advanced signal processing are widely studied to enhance imaging capability and adjustability. Heterogeneous image fusion is a typical image-enhancing technique that combines at least two distinct source or sensor images into a single composite image. The fused image will be more informative, which may be deteriorated and limited by low signal-to-noise ratio condition. Hence, many applications are benefited from computational theory of multi-source image fusion, especially for the social security surveillance, network security monitoring, network video analysis, and military observation.<sup>1-3</sup>

According to the heat radiation emitted by the object, infrared rays can be used to detect hidden heat sources in complex environments. To capture the heat radiation of specific targets, the infrared image sensor is

designed in such a way that it is only sensitive to light wavelength between 780 and 2000 nm, but at the cost to lose the information of visible light parts. On the contrary, visible images contain more rich texture of surrounding environment, which may also be important to effective target observation.<sup>4,5</sup> In this article, we discuss the visible and infrared image fusion problem by contrast and directional features optimization (CDFO).

The image fusion can usually be divided into three levels: pixel level, feature level, and decision level. Traditional feature-level fusion method includes

<sup>1</sup>School of Computer Science and Engineering, Xi'an University of Technology, Xi'an, China

<sup>2</sup>Shaanxi Key Laboratory for Network Computing and Security Technology, Xi'an, China

## Corresponding author:

Haiyan Jin, School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China.

Email: jinhaiyan@xaut.edu.cn



weight-averaging method, Bayesian estimation method, and cluster analysis method.<sup>6</sup> These methods combine the features extracted from multi-source information on different sensors (including shape, edge, region, contour, texture, and horn). The feature extraction and comprehensive analysis can be well processed even with number of different sensor images. However, the hierarchical information fusion at different resolutions is still an open problem. Decision-level fusion method<sup>7</sup> could find feature correspondence via classification and identification of each image, and then further fusion process will be carried out by global optimization. The non-uniform resolution problem cannot produce a good fusion result due to simple decision-level fusion with multi-resolution images. In recent years, pixel level-based fusion methods<sup>8,9</sup> show more advantages on those cases with complex images. For example, Kannan and Perumal<sup>10</sup> proposed a pixel-level image fusion method based on discrete wavelet transform (DWT), in which an image can be decomposed into a sequence of different spatial resolution images. Then, the fusion process can be applied on different image scales. Li and Dong<sup>11</sup> and He et al.<sup>12</sup> presented a contrast pyramid (CP) fusion method. The original image is resolved into a series of decomposing layers with different resolution and frequencies. The fusion process is carried out separately on the components of each spatial frequency layer. In order to get better direction information, Jin et al.<sup>13,14</sup> suggested fusing the satellite infrared image with visible image by introducing directional filter banks. This method can extract the directional sub-graph of infrared image, so that the directional information can be well preserved during the fusion process.

The signal processing by human visual system is performed on different channels of different scales,<sup>15</sup> so we introduced CP to pre-processing. Considering the direction information in full and the edge feature of high-frequency information, directional filter banks are used to get more direction features. First, the CP is used to decompose the image and adjust the contrast in different scales. The input multi-channel images are decomposed on different frequency bands, and the multi-resolution CP sequences form a pyramid-like structure. Each scale of CP corresponds to different spatial frequency characteristics. Second, we use directional filter banks to process the decomposition image on the basis of CP. Filter group provides an efficient analysis and synthesis of discrete signal structure. Specific structure of filter bank can obtain continuous multi-resolution base. Hence, we can obtain the abundant direction information and multi-scale information. During the fusion process, traditional methods determine the fusion coefficient weights according to the experience of human, but we cannot get a good fusion effect. Therefore, whale optimization algorithm (WOA) is introduced to search the optimal fusion coefficients adaptively. Compared to majority of the traditional methods, the proposed method based on CDFO provides an accurate image fusion result for getting the most significant features from the multiple sensor images. In this article, the overall technical route of image fusion based on CDFO is shown in Figure 1.

## Related work

Simple pixel-level fusion can be treated as a linear summation of pixels from the distinct source images,<sup>16</sup> as shown in equation (1)

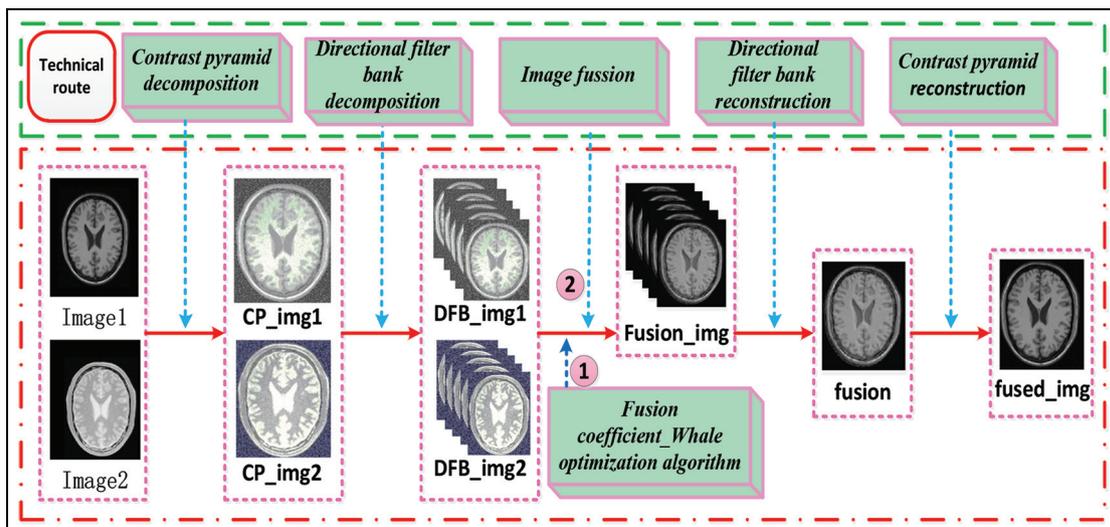
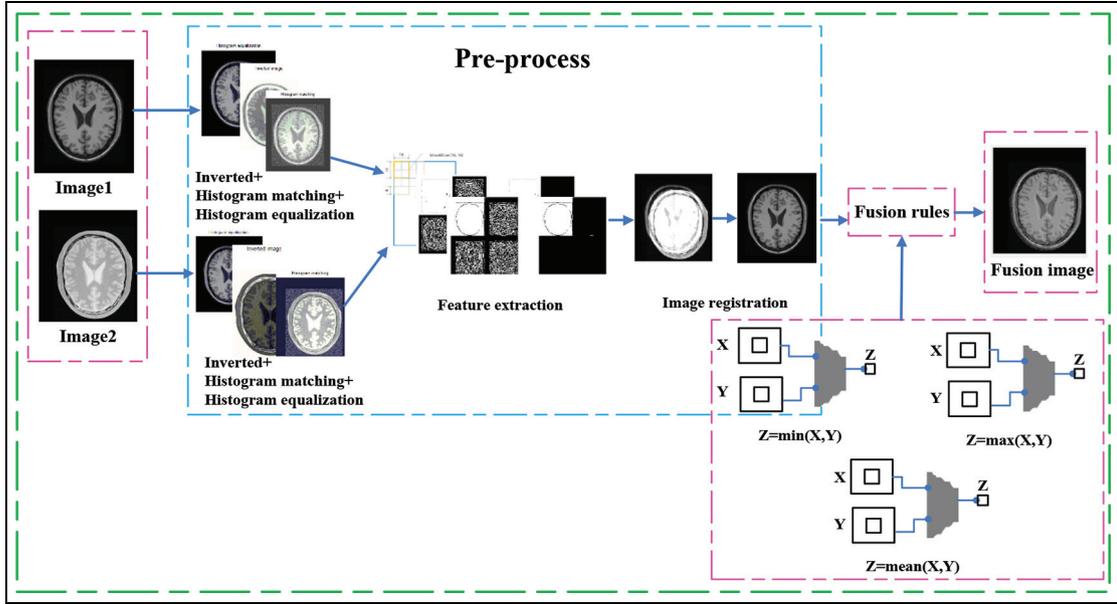


Figure 1. Overall technical route of CDFO image fusion.



**Figure 2.** Pixel-level image fusion mechanism.

$$g(x, y) = a_1 f_1(x, y) + a_2 f_2(x, y) + \dots + a_n f_n(x, y) \quad (1)$$

where  $a_1 + a_2 + \dots + a_n = 1$  represents the total contributions of different sources that should be normalized for the fusion image result  $g(x, y)$ . The mechanism of pixel-level image fusion is shown in Figure 2. For two original images, pre-treatment is done first: inverted histogram matching and histogram equalization are performed; feature extraction and image registration are processed on it; next, fusion rule is applied to fuse image; finally, the corresponding coefficients are selected through the fusion rules to get the final fused image.

### CP decomposition

In order to obtain multi-resolution-level image information, the original visible and infrared image should be pre-processed. The principles of CP decomposition are as follows.

**Establishing Gaussian decomposition.** CP is based on Gaussian pyramid. First, we use  $I(i, j)$  to represent the original image. Second, Gaussian pyramid<sup>17,18</sup> is used to get the decomposition image  $G_l$ . Finally, interpolation method is used to enlarge the  $G_l$ , so that we can get the image  $G_l^*$ . The size of  $G_l^*$  is the same as  $G_{l-1}$ . The equation is as follows

$$G_l^*(i, j) = 4 \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) G_l' \left( \frac{i+m}{2}, \frac{j+n}{2} \right) \quad (2)$$

where  $0 < l \leq N$ ,  $0 < i \leq u$ , and  $0 < j \leq v$ ; and  $u$  and  $v$  are the number of rows and columns, respectively. In the  $l$  layer,  $w(m, n)$  is a window function and is actually a low-pass filter that satisfies the following constraints

1. *Separability*:  $w(m, n) = w(m) \cdot w(n)$ ,  $m \in [-2, 2]$ ,  $n \in [-2, 2]$
2. *Normalization*:  $\sum_{n=-2}^2 w(n) = 1$
3. *Symmetry*:  $w(n) = w(-n)$

The limitation of the above constraint is to ensure the low-pass property and the smoothness of the image after the image is reduced and expanded, and no seam effect occurs. Thus, we construct  $w(0) = 3/8$ ,  $w(1) = w(-1) = 1/4$ , and  $w(2) = w(-2) = 1/16$  and use the common window width of  $5 \times 5$ .

**Building the CP.** In equation (2), we can calculate  $G_l'$  as follows

$$G_l' \left( \frac{i+m}{2}, \frac{j+n}{2} \right) = \begin{cases} G_l \left( \frac{i+m}{2}, \frac{j+n}{2} \right), & \frac{i+m}{2}, \frac{j+n}{2} \text{ is integer} \\ 0, & \text{others} \end{cases} \quad (3)$$

The window function has low-pass filter properties, thus  $G_l^*$  can be seen as the background of  $G_l$ . Therefore, the definition of CP can be decomposed as equation (4)

$$C_l = \frac{G_l}{G_{l+1}} = \frac{G_l}{\text{Expand}(G_{l+1})} - I, \quad 0 \leq l < N \quad (4)$$

$$C_N = G_N, \quad l = N$$

where  $C_l$  is the  $l$ th layer of CP decomposition;  $G_l$  is the  $l$ th layer of Gaussian decomposition; and  $\text{Expand}(G_{l+1})$  is the interpolation magnification of  $G_{l+1}$ , and its size is same as  $G_l$ .

**Rebuilding the original image from the CP.** Transforming equation (4), we can get a reconstruction equation of CP

$$G_N = C_N, \quad l = N$$

$$G_l = (C_l + I)\text{Expand}(G_{l+1}), \quad 0 \leq l < N \quad (5)$$

According to equation (5), we can iteratively start from the top layer,  $C_N$ , of CP ( $C_N, C_{N-1}, \dots, C_0$ ) and make  $l = N, N-1, \dots, 0$  at the same time. Each layer of the Gaussian pyramid can be obtained. Finally, the precise reconstruction of the original image,  $G_0$ , can be recovered exactly by reversing the above steps.

### Construction of directional features

Considering the multi-sensor image characteristics, we use directional filter banks<sup>19,20</sup> to capture the directional features. The traditional directional filter banks lack multi-scale analysis capabilities, so in this article, we design a new directional filter bank to make it suitable for multi-sensor visible and infrared image.

It consists of two-channel diamond filter bank and parallelogram filter bank that cascade through the tree structure and through the  $L$ -level binary tree decomposition; the spectrum plane  $[-\pi, \pi]$  is decomposed into  $2^l$  sub-bands and the band is divided into wedges, where each wedge spectral sub-band extracts image information that is perpendicular to its direction, where  $l$  is the cascade series as shown in Figure 3.

In Figure 3(b),  $Q_1, D_0, D_1, D_2,$  and  $D_3$  are the inseparable sampling matrices, as shown in equation (6)

$$Q_1 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad D_0 = \begin{pmatrix} 0 & 2 \\ -1 & 1 \end{pmatrix},$$

$$D_1 = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, \quad D_3 = \begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix} \quad (6)$$

Figure 3(b) shows the three-stage cascaded decomposition of the directional filter bank (the reconstructed part is the inverse of the decomposed part). The wedge-shaped directional spectrum partitioning makes it suitable for extracting the edge contours with spatial localization in the image.

### The WOA

Traditional fusion uses the maximum fusion rule or experience to determine the fusion coefficients, thus the fusion effect can be greatly deteriorated by unreasonable coefficient selection. In this article, we adopt a heuristic intelligent optimization algorithm to determine the optimal fusion coefficients.

Some brain regions of the whale are similar to the brain structure of human beings.<sup>21</sup> Prey is one of the most characteristic behaviors of the whale. By studying the hunting behaviors of whales' group in the nature environment, researchers have proposed a meta-heuristic optimization algorithm which is called WOA. In 2015, Mirjalili and Lewis<sup>22</sup> proposed this method by simulating the hierarchy of the whale population and their hunting behaviors.

### The proposed CDFO method

#### Contrast and directional features

The traditional CP decomposition has good physical meaning but ignores the direction information and affects the fusion effect. The directional filter bank adds direction based on the CP decomposition, which not only maintains the physical meaning of contrast but also detects the direction information well and can provide more feature information for the fused image.

In directional filter banks, we define a sampling of multi-dimensional space  $Z^d$  in the grid, and the grid is expressed by a non-singular integer matrix  $M$ . The size of  $M$  is  $d \times d$ . The equation of  $M$  is as follows

$$\text{LAT}(M) = \{M_n, n \in Z^d\} \quad (7)$$

First, we give a proposition:  $\text{LAT}(A) = \text{LAT}(B)$ , if and only if  $A = BE$ , among them,  $E$  is an unimodular integer matrix.

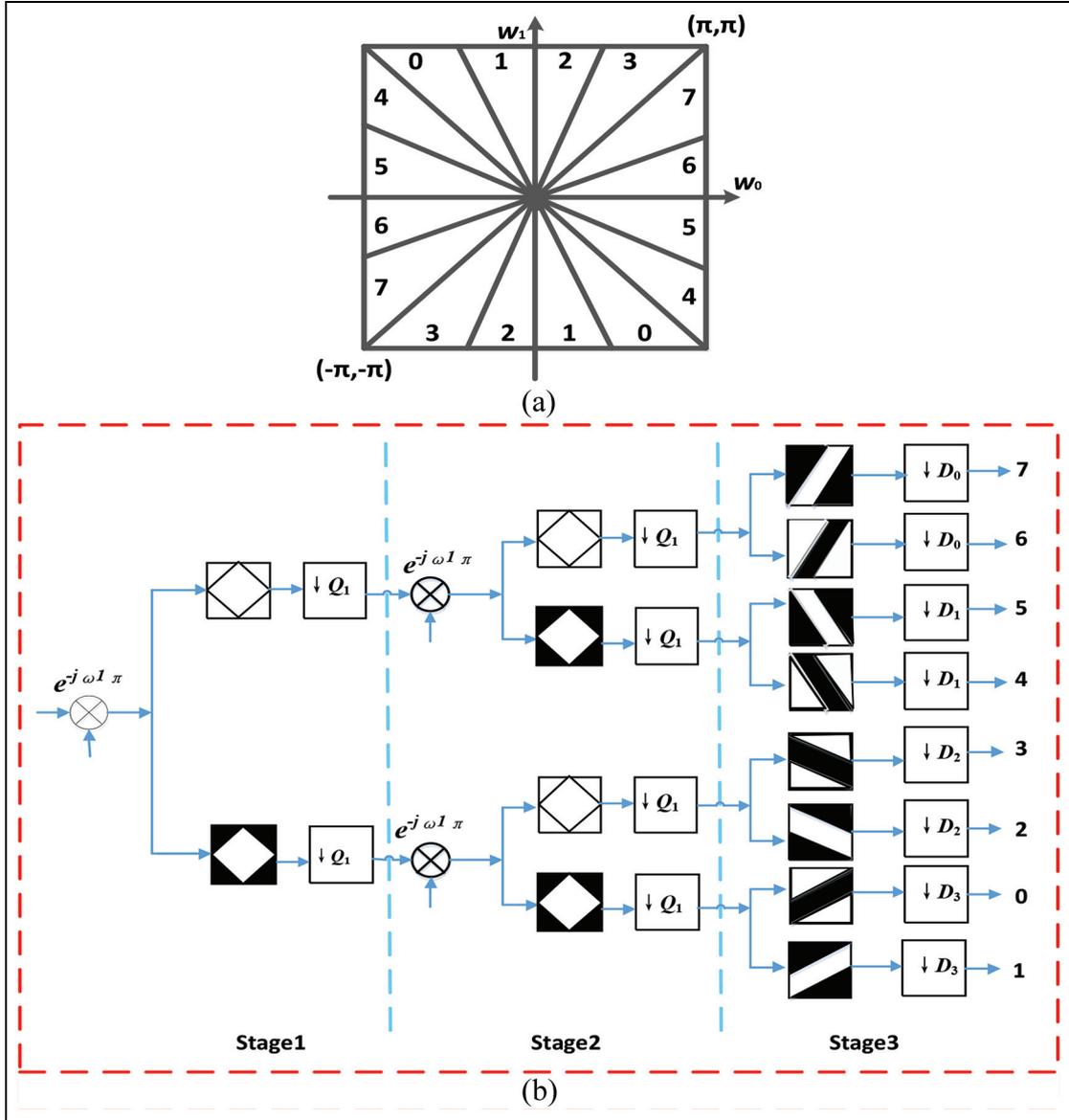
In this article, we use the decomposition of two-dimensional, two-channel ladder structure filter and quincunx sampling network. Two matrices are used to represent the quincunx sub-network

$$Q_0 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad Q_1 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad (8)$$

In directional filter banks, the following four basic unimodular matrices are used to achieve the invariance of rotation operation

$$R_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad R_1 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix},$$

$$R_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad R_3 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \quad (9)$$



**Figure 3.** Directional filter bank spectral segmentation: (a) spectrum division diagram and (b) decomposition part.

here, we use Smith's decomposition which can diagonalize any integer matrix  $M$  into a product form of  $UDV$ . In this form,  $U$  and  $V$  are the unimodular integer matrices and  $D$  is an integer diagonal matrix.<sup>23</sup> Thus, the quincunx matrix in equation (8) can be represented by Smith's form as follows

$$\begin{aligned} Q_0 &= R_1 D_0 R_2 = R_2 D_1 R_1 \\ Q_1 &= R_0 D_0 R_3 = R_3 D_1 R_0 \end{aligned} \quad (10)$$

where  $D_0 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$  and  $D_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ .  $D_0$  and  $D_1$  are the two-dimensional diagonal matrix and each dimension corresponds to a binary sampling.

In the first two layers of directional filter banks, quincunx filter banks (QFBs) will be used on each

layer.  $Q_0$  is the sampling matrix on the first layer and  $Q_1$  on the second layer. From the third layer, the rest of the tree structure begins to extend. To obtain better frequency division, we use the QFB for resampling. The first half of the channel uses  $R_0$  and  $R_1$  resampling and gets a sub-band based on the direction of the basic level:  $+45^\circ$  and  $45^\circ$ . The second half of the channel is resampled with  $R_2$  and  $R_3$  to produce the remaining sub-bands.

We can simplify the total sampling matrix of the QFB with the quincunx matrix of Smith's form. We can calculate the sampling matrix using the following equation

$$\begin{aligned} p_0 &= R_0 Q_0 = D_0 R_2 \\ p_1 &= R_1 Q_1 = D_0 R_3 \end{aligned} \quad (11)$$

According to the proposition presented above, we can get the following equation

$$LAT(p_0) = LAT(p_1) + LAT(D_0) \quad (12)$$

The above equation shows the sampling network of QFB, which uses the  $R_0$  and  $R_1$  resampling and is equivalent to the down-sampling that is along the  $n_0$ -dimensional space. Although CP with directional filter banks has some excellent features of multi-scale and multi-directional geometric analysis, for image fusion algorithms, the choice of fusion coefficients is also very important. WOA has attracted much more attention due to its excellent performance and powerful search capability.

### Fusion coefficient optimization with the WOA

The traditional fusion coefficient is selected through human experience. In this article, WOA is introduced to optimize the adaptive fusion coefficient, so that the image contrast information and direction information can be maximally retained and the fusion effect is obviously improved.

The WOA builds a global search mechanism by simulating the hunting behavior of whale, which can simulate the attacking mechanism of humpback whales (bubble-net hunting). The details of the WOA algorithm are described as follows.

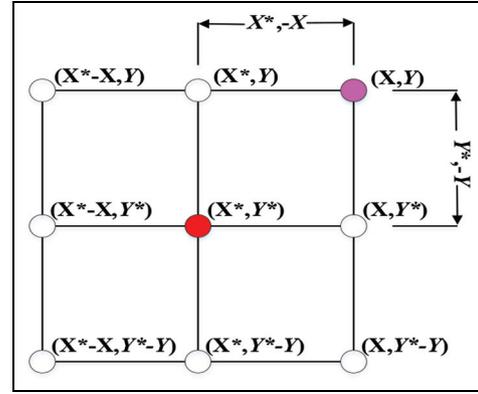
**Encircling prey.** Before hunting, humpback whales identify the location of prey and then encircle their prey. In search space, the optimal position of the target prey is unknown at the beginning of time, thus WOA assumes that the current optimal solution is the target prey or close to the optimal position. When we ascertain the optimal search agent, other searches will update their own position based on this agent. Mirjalili and Lewis<sup>22</sup> described this behavior in a mathematical model as follows

$$\begin{aligned} \vec{D} &= |\vec{C} \cdot \vec{X}^*(t) - \vec{X}(t)| \\ \vec{X}(t+1) &= \vec{X}^*(t) - \vec{A} \cdot \vec{D} \end{aligned} \quad (13)$$

where  $t$  indicates the current iteration,  $\vec{A}$  and  $\vec{C}$  are the coefficient vectors,  $X^*$  is the position vector of the target, and  $\vec{X}$  is the optimal position vector of the humpback whales. We can calculate  $\vec{A}$  and  $\vec{C}$  using the following equation

$$\begin{aligned} \vec{A} &= 2\vec{a} \cdot \vec{r} - \vec{a} \\ \vec{C} &= 2 \cdot \vec{r} \end{aligned} \quad (14)$$

here, the value of  $\vec{a}$  is linearly decreased from 2 to 0 and  $\vec{r}$  is the random vector in the interval  $[0, 1]$ .



**Figure 4.** Predicted possible locations of position vectors in 2D space.

The two-dimensional feedback behavior is shown in Figure 4. In Figure 4,  $X^*$  is the best position we have got so far, and we can update the current position  $(X, Y)$  by  $(X^*, Y^*)$ . The mathematical model to describe the hunting behavior of bubble-net is mentioned in equation (13).

Bubble-net searching mechanism has two patterns: the first is shrinking encircling pattern and the second is spiral updating position.

1. *Shrinking encircling pattern.* Like a quadrangle, we updated the location of the optimal search agent by changing the value of  $\vec{a}$  in equation (14), thus the fluctuation range of  $\vec{A}$  decreased. When the value of  $\vec{a}$  decreased from 2 to 0,  $\vec{A}$  becomes a random value between  $-a$  and  $a$ ; When the value of  $\vec{A}$  falls between  $-1$  and  $1$ , the next position of humpback whales can be any position between their current position and the prey. In Figure 5, we can observe this relationship clearly.
2. *Spiral updating position.* In this mathematical model, we use a spiral equation to calculate the distance between whale  $(X, Y)$  and prey  $(X^*, Y^*)$  as follows

$$\vec{X}(t+1) = \vec{X}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) \quad (15)$$

where  $\vec{D}' = |\vec{X}^*(t) - \vec{X}(t)|$ , in which  $\vec{D}'$  indicates the distance between the  $i$ th whale and the prey;  $b$  is a constant for defining the shape of the spiral equation; and  $l$  is a random number in the interval  $[-1, 1]$ .

In the process of optimization, we assume that if  $p < 0.5$ , we use the mechanism of shrinking encircling pattern; and if  $p \geq 0.5$ , the spiral updating position is applied. The mathematical model is as follows

$$\vec{X}(t+1) = \begin{cases} \vec{X}^*(t) - \vec{A} \cdot \vec{D}, & \text{if } p < 0.5 \\ \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t), & \text{if } p \geq 0.5 \end{cases} \quad (16)$$

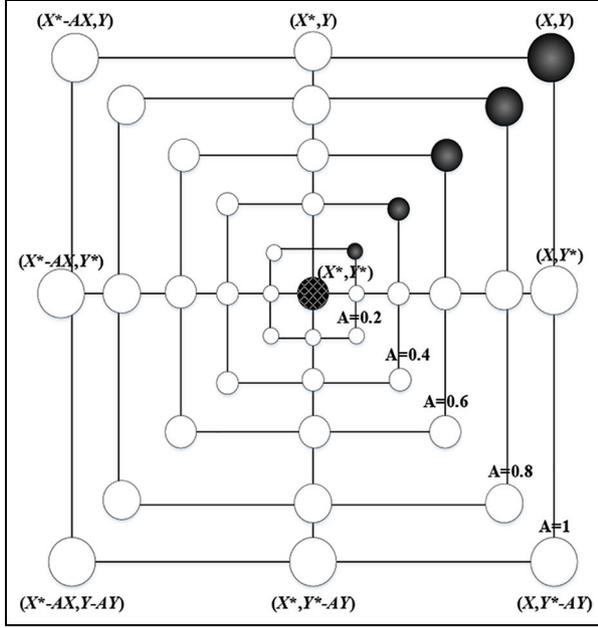


Figure 5. Shrinking encircling mechanism.

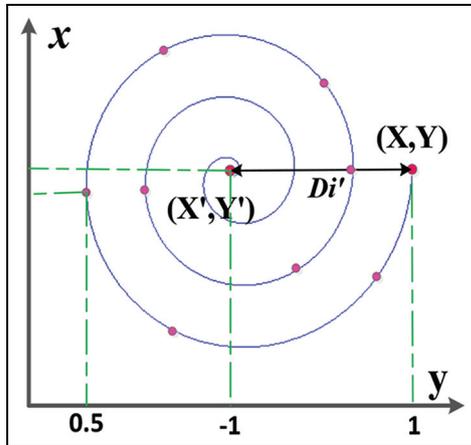


Figure 6. Mechanism of bubble-net searching with spiral pattern.

where  $p$  is a random number in the interval  $[0, 1]$ . Figure 6 shows the spiral structure clearly.

**Search for prey (exploration phase).** Unlike the exploitation phase, the search agent updates its location according to the random selection during the exploration phase, thus the random value of  $\vec{A}$  is used to force the search agent to stay away from the reference whale, and in this situation, the value of  $\vec{A}$  is greater than 1 or less than  $-1$ . The mathematical model is as follows

$$\begin{aligned} \vec{D} &= |\vec{C} \cdot \vec{X}_{rand} - \vec{X}| \\ \vec{X}(t+1) &= \vec{X}_{rand} - \vec{A} \cdot \vec{D} \end{aligned} \quad (17)$$

#### Algorithm 1. Whale optimization algorithm

1. Initialize the whale population,  $X_i$  ( $i = 1, 2, \dots, n$ );
2. Calculate the fitness of each search agent;
3.  $X^*$  = the best search agent;
4. while ( $t <$  maximum number of iterations)
5. for each search agent
6. Update  $a$ ,  $A$ ,  $C$ ,  $l$ , and  $p$ ;
7. if ( $p < 0.5$ )
8. if ( $|A| < 1$ )
9. Update the position of the current search agent by equation (13);
10. else
11. Select a random search agent ( $X_{rand}$ );
12. Update the position of the current search agent by equation (17);
13. end if
14. else
15. Update the position of the current search agent by the equation (15);
16. end if
17. end for
18. Check if any search agent goes beyond the search space and amend it;
19. Calculate the fitness of each search agent;
20. Update  $X^*$  if there is a better solution;
21.  $t = t + 1$ ;
22. end while
23. Return  $X^*$ ;

where  $p_i$ ,  $p = \{p_1, p_2, \dots, p_n\}$ , is a random position vector chosen from the current population.

On the whole, WOA algorithm is a global optimization, because it has the capabilities of exploration and exploitation. Pseudo code of WOA is given as follows.

To generalize the proposed algorithm, our fusion can be summarized by the following steps:

1. Initialize parameters. We adopt window function,  $w$ , to decompose, and the decomposition level is set as 3. The ‘‘Phoong, Kim, Vaidyanathan and Ansari’’ directional filter banks of the trapezoidal structure are used, and the maximum decomposition layer is 4.
2. Decompose CP on each original image separately and obtain the decomposition images  $M_1$  and  $M_2$  with the size of  $1/2^3$  of the original image.
3. Use the directional filter banks built according to section ‘‘Fusion coefficient optimization with the WOA,’’ and decompose  $M_1$  and  $M_2$  to get the filtering images  $M'_1$  and  $M'_2$  and figures  $D_1$  and  $D_2$  carried by directions at this layer.
4. Take the filtering images  $M'_1$  and  $M'_2$  obtained in equation (3) as inputs and iterate and implement step 3 until the completion of the set decomposition layer.
5. In each final obtained image, add information  $D_1$  and  $D_2$  carried by directions into separate

decomposition results  $y_1$  and  $y_2$  as the row vectors.

6. Conduct the low-pass filtering of decomposition results  $y_1$  and  $y_2$  separately and obtain two groups of decomposition coefficients  $C_1$  and  $C_2$ .
7. Use the WOA proposed in this article to optimize the coefficients  $C_1$  and  $C_2$  to obtain the new fusion coefficient  $C$ .
8. Consider  $C$  as the input coefficient of directional filter banks for reconstitution to obtain reconstitution result  $y$ .
9. Conduct contrast reconstruction and obtain the final fusion result  $Y$ .

### Computational complexity analysis

The computational complexity of the algorithm in this article consists of two parts: the computational complexity of the algorithm in section “Contrast and directional features” and that of the algorithm in section “Fusion coefficient optimization with the WOA.”

In section “Contrast and directional features,” CP decomposes the original images first before entering the directional filter bank. After down-sampling and filtering, the amount of data is reduced greatly, thereby improving the efficiency. Suppose the size of both images to be fused is  $N \times N$ , and the number of levels to be CP decomposed is  $l$ , then after down-sampling and filtering, the size of images is  $(N/2^l) \times (N/2^l)$ . Hence, the size of the images to enter the directional filter bank is  $(N/2^l) \times (N/2^l)$ . The element number contained in two groups of coefficients,  $C_1$  and  $C_2$ , can be calculated by formula (18)

$$M = \frac{5}{4}(N/2^l)^2 = \frac{5N^2}{4^{l+1}} \quad (18)$$

where  $N$  is the size of the images to be fused,  $l$  is the number of levels to be CP decomposed, and  $M$  is the element number contained in two groups of coefficients, “ $C_1$ ” and “ $C_2$ .” The complexity of comparison operation for finding the maximal absolute value is  $O(M)$ , that is,  $O(5N^2/4^{l+1})$ , which is the computational complexity of the algorithm in section “Contrast and directional features.”

In the optimization of the fusion coefficient in section “Fusion coefficient optimization with the WOA,” if the maximum number of iterations is  $k$  and the size of the whale population is  $N$ , then the time complexity of the optimization coefficient of the WOA optimization algorithm is  $O(kN)$ .

As a whole, the computational complexity of the algorithm in this article is  $O\left(\frac{5N^2}{4^{l+1}}\right) + O(kN)$ , which can be regarded as  $O(5N^2/4^{l+1})$ .

## Experiment and analysis

### Numerical indexes

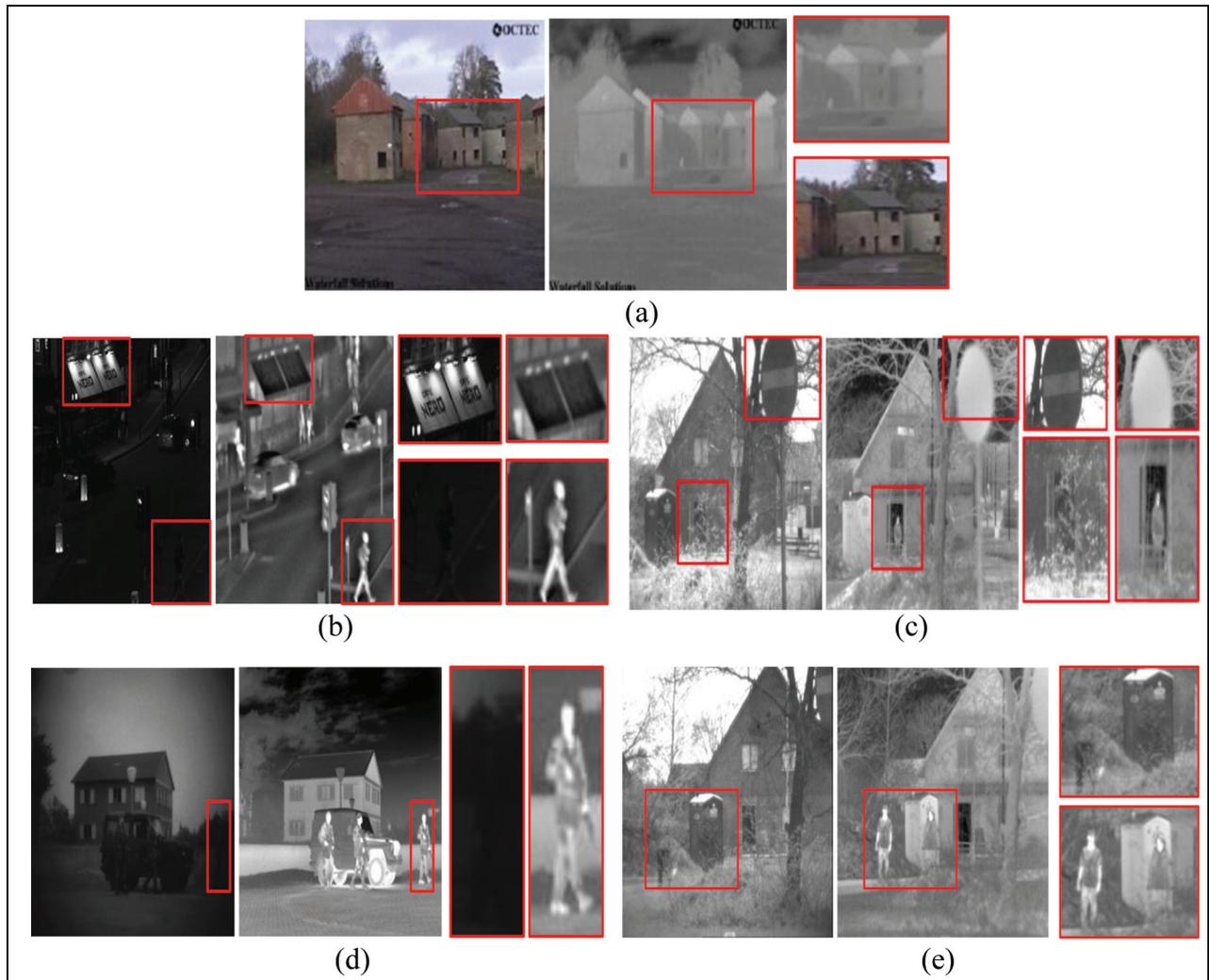
At present, there is no unified standard to evaluate the fusion effect for all kinds of images. In this article, we adopt some evaluation metrics presented in relevant references to evaluate the fusion results objectively, such as information entropy (IE), average grads, standard deviation, and spatial frequency. For example, the IE of the image reflects the amount of information of the image, where the larger the value, the more information the image contains. The average gradient (AG) of the image reflects the ability of the image to express contrast to the tiny details. The standard deviation (SD) of the image is a measure of the degree of dispersion of the pixel values of the image compared to the pixel average, which is the contrast information of the image, where the larger the value, the greater the contrast of the image. The spatial frequency (SF) effectively reflects the details of the image, where the larger the value, the clearer the image. The edge strength (ES) value of the image is an evaluation index that describes the visually important edge intensity information and direction information of the fusion result. The execution time (ET) reflects the efficiency of different fusion methods.

### Experimental analysis

In order to verify the advantages of the proposed algorithm, we choose several fusion algorithms to compare, such as fast filtering image fusion (FFIF) method,<sup>24</sup> CP-based fusion method,<sup>11,12</sup> DWT-based fusion method,<sup>10</sup> and principal component analysis (PCA)-based fusion method.<sup>25</sup> As shown in Figure 7, the first dataset represents a section ( $256 \times 256$  pixels) of visible and infrared images without smoke shielding acquired by a Sony Camcorder and a long-wave infrared (LWIR) sensor. The second group of visible and infrared images is near the street signs and a man in the doorway. The third group of images ( $256 \times 256$  pixels) is crossroads at night. The fourth group of images ( $256 \times 256$  pixels) is three soldiers and a jeep, and the last group of images ( $256 \times 256$  pixels) is two men in front of a house. The fusion results are shown in Figure 8.

In the experiments, fusion performance is evaluated by qualitative analysis and objective indicator (as shown in Tables 1–5). Qualitative analysis is adopted by the visual effect and spectral fidelity, including spatial resolution, clarity, and sophisticated details, which indicate the extent of preserving original spectral signal or characteristics.

From the visual effects (Figure 8), the overall profile of the FFIF fusion result is very rough and the saturation is distorted, especially in the second group of experiments; CP-based fusion results have higher



**Figure 7.** Five pairs of visible and infrared images: (a) building images without smoke shielding, (b) man in the doorway, (c) crossroads at night, (d) soldiers with jeep, and (e) two men in front of a house. Left column: visible image; middle column: infrared image; right column: partial enlargement image.

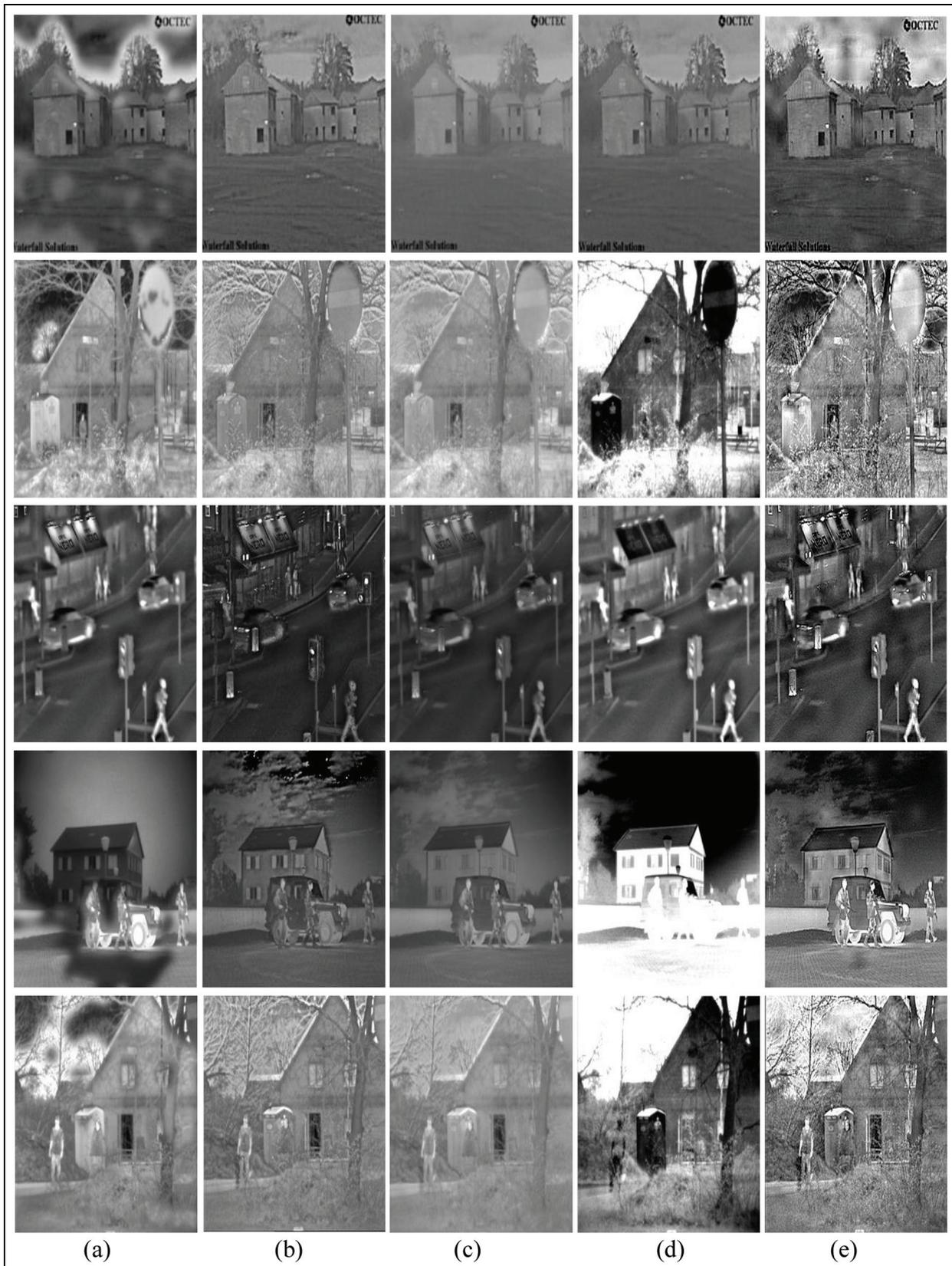
contrast, but the edge contour is not clear enough; DWT-based fusion results lack spectral characteristics, and the sharpness is not enough; PCA-based fusion images have excessive brightness, rough edge contours, loss of a large amount of detail information, and severe distortion of saturation; the proposed CDFO method has better fusion results than the previous four groups of experiments. Rich-structure features and clear-edge contours can better highlight the target area of interest.

By the numerical results (Tables 1–5), the evaluation indexes of the proposed CDFO method are mainly the highest, which indicate that the fusion images have the highest sharpness and the most abundant details. However, in terms of the SD indicators, the SD values of the CP-based fusion results in the second, fourth, and fifth groups are higher. The reason is that the contrast of the fusion results based on the CP is too large, and its visual effect is particularly poor. Also, the time

cost of the CDFO method is almost lower, which is very important in practical applications.

The convergence curve of the fusion coefficient optimization of the WOA in five different groups of images is shown in Figure 9. The second and third groups of images tend to converge when iterated 4 times, and the fusion result no longer changes. The first and fifth sets of images also reached a steady state when iterated 6 times, and the fourth set of images also converge when the number of iterations reached 8. Therefore, the figure shows that the image fusion method based on WOA for coefficient optimization has a fast convergence speed and good convergence stability, which can effectively save the optimization time of the fusion coefficient.

In summary, the above five sets of simulation experiments have proved that both on the subjective vision and the quantitative numerical indicators, also taking



**Figure 8.** Fusion results by different methods: (a) FFIF-based, (b) CP-based, (c) DWT-based, (d) PCA-based, and (e) CDFO-based.

**Table 1.** Numerical results of the first line of Figure 8.

Methods	IE	AG	SD	SF	ES	ET
IR	6.3274	2.0478	23.0849	5.4848	2.1069	–
VS	7.0887	4.1367	61.0146	10.5986	4.5757	–
FFIF	6.9156	4.5399	36.7987	10.6207	4.9612	0.7513
CP	6.4694	4.7274	27.4743	10.5398	5.1311	0.7643
DWT	6.248	2.4895	15.5729	6.5587	2.5496	1.0064
PCA	7.1936	5.0315	70.3889	12.5029	5.6660	<b>0.4939</b>
CDFO	<b>7.2640</b>	<b>11.2167</b>	<b>49.1775</b>	<b>25.5808</b>	<b>12.3569</b>	0.5086

IE: information entropy; AG: average gradient; SD: standard deviation; SF: spatial frequency; ES: edge strength; ET: execution time; IR: original infrared image; FFIF: fast filtering image fusion; CP: contrast pyramid; DWT: discrete wavelet transform; PCA: principal component analysis; CDFO: contrast and directional features optimization; VS: original visible image.

Significant of bold values represents the optimal numerical value gained from different fusion methods using the corresponding evaluation index.

**Table 2.** Numerical results of the second line of Figure 8.

Methods	IE	AG	SD	SF	ES	ET
IR	7.1771	6.7654	31.3187	13.3196	8.4907	–
VS	7.2515	10.7926	58.4513	23.6716	10.6270	–
FFIF	7.484	10.3536	44.9276	21.2707	10.4554	0.8217
CP	7.0869	13.0236	34.0332	26.0642	15.2924	0.7701
DWT	6.7888	7.0359	27.167	13.7803	7.6928	0.9890
PCA	6.6373	14.4227	<b>85.9935</b>	33.9517	9.6831	0.5097
CDFO	<b>7.2640</b>	<b>11.2167</b>	49.1775	<b>25.5808</b>	<b>12.3569</b>	<b>0.4322</b>

IE: information entropy; AG: average gradient; SD: standard deviation; SF: spatial frequency; ES: edge strength; ET: execution time; IR: original infrared image; FFIF: fast filtering image fusion; CP: contrast pyramid; DWT: discrete wavelet transform; PCA: principal component analysis; CDFO: contrast and directional features optimization; VS: original visible image.

Significant of bold values represents the optimal numerical value gained from different fusion methods using the corresponding evaluation index.

**Table 3.** Numerical results of the third line of Figure 8.

Methods	IE	AG	SD	SF	ES	ET
IR	6.7243	5.3714	32.785	12.2326	6.0719	–
VS	5.0389	3.521	27.009	13.8142	6.9118	–
FFIF	6.7468	6.0716	34.7583	14.7239	8.7572	0.8790
CP	6.0738	7.8986	25.1093	20.9657	12.3687	0.7634
DWT	6.1205	4.075	23.0953	10.0464	5.5955	0.9884
PCA	6.9782	6.4662	38.7415	14.5102	7.6706	0.4948
CDFO	<b>6.9771</b>	<b>16.1316</b>	<b>42.3714</b>	<b>37.2714</b>	<b>22.7072</b>	<b>0.4870</b>

IE: information entropy; AG: average gradient; SD: standard deviation; SF: spatial frequency; ES: edge strength; ET: execution time; IR: original infrared image; FFIF: fast filtering image fusion; CP: contrast pyramid; DWT: discrete wavelet transform; PCA: principal component analysis; CDFO: contrast and directional features optimization; VS: original visible image.

Significant of bold values represents the optimal numerical value gained from different fusion methods using the corresponding evaluation index.

**Table 4.** Numerical results of the fourth line of Figure 8.

Methods	IE	AG	SD	SF	ES	ET
IR	7.528	4.9375	57.8926	12.6010	3.6485	–
VS	6.753	2.1102	38.7895	4.9764	2.4612	–
FFIF	7.2801	4.2296	44.2637	10.6284	2.5793	0.7977
CP	6.5778	6.1498	28.0864	14.5947	6.5410	0.7591
DWT	6.6672	3.2526	25.8714	7.6963	2.5991	0.9833
PCA	4.7511	5.1154	<b>110.5146</b>	17.6021	5.1737	0.5181
CDFO	<b>7.6542</b>	<b>15.4464</b>	58.4816	<b>32.7404</b>	<b>12.2877</b>	<b>0.5005</b>

IE: information entropy; AG: average gradient; SD: standard deviation; SF: spatial frequency; ES: edge strength; ET: execution time; IR: original infrared image; FFIF: fast filtering image fusion; CP: contrast pyramid; DWT: discrete wavelet transform; PCA: principal component analysis; CDFO: contrast and directional features optimization; VS: original visible image.

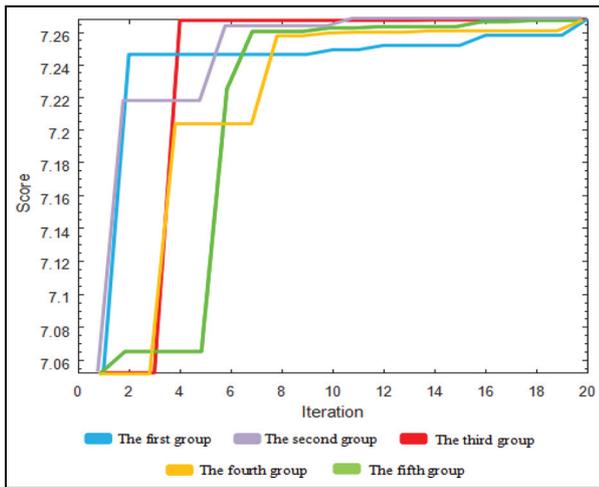
Significant of bold values represents the optimal numerical value gained from different fusion methods using the corresponding evaluation index.

**Table 5.** Numerical results of the fifth line of Figure 8.

Methods	IE	AG	SD	SF	ES	ET
IR	7.0503	5.2079	32.5811	10.0206	5.9683	–
VS	6.9399	7.8798	57.1193	18.7600	10.7481	–
FFIF	7.2079	7.5527	39.0904	16.5521	10.8030	0.7638
CP	6.6873	9.7165	27.835	20.3017	12.9452	0.7973
DWT	6.4732	5.2818	23.0406	10.8628	7.1504	1.0566
PCA	6.9152	10.7111	<b>79.5134</b>	25.4203	7.4961	0.4832
CDFO	<b>7.4034</b>	<b>27.5327</b>	55.493	<b>54.0799</b>	<b>36.1802</b>	<b>0.4404</b>

IE: information entropy; AG: average gradient; SD: standard deviation; SF: spatial frequency; ES: edge strength; ET: execution time; IR: original infrared image; FFIF: fast filtering image fusion; CP: contrast pyramid; DWT: discrete wavelet transform; PCA: principal component analysis; CDFO: contrast and directional features optimization; VS: original visible image.

Significant of bold values represents the optimal numerical value gained from different fusion methods using the corresponding evaluation index.

**Figure 9.** Convergence curves of CDFO method applied in five different sets of images.

into account the time cost, the proposed CDFO method has obvious advantages.

## Conclusion

In this article, a CDFO image fusion algorithm is proposed to take advantages of both infrared images and visible images, which can better reconstruct the target scene using CDFO. The method can preserve more structural features and color spectrum information. At the same time, considering the high-frequency information with different scales, we construct directional filter banks to capture directional features. Finally, the WOA is applied to optimize the fusion coefficient. The proposed method has great advantages in both image quality and quantity. By testing on real visible and infrared image datasets, the proposed algorithm can preserve more scene details and information, which is more valuable for video monitoring and security surveillance in low visible environment.

## Acknowledgements

H.J. and Z.X. contributed to conceptualization, methodology, and writing review and editing; M.Z. and Y.L. contributed to software, validation, formal analysis, and writing of original draft preparation; and H.J. contributed to supervision, project administration, and funding acquisition.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported in part by the National Natural Science Foundation of China under grant no. 61472204.

## References

1. Xydeas CS and Petrovic V. Objective image fusion performance measure. *Electron Lett* 2000; 36(4): 308–309.
2. Quan S, Qian W, Guo J, et al. Visible and infrared image fusion based on curvelet transform. In: *Proceedings of the 2nd international conference on systems and informatics (ICSAI)*, Shanghai, China, 15–17 November 2014, pp.828–832. New York: IEEE.
3. Fu Z, Dai X, Li Y, et al. An improved visible and infrared image fusion based on local energy and fuzzy logic. In: *Proceedings of the 12th international conference on signal processing (ICSP)*, Hangzhou, China, 19–23 October 2014, pp.861–865. New York: IEEE.
4. Fei C, Zhang P, Tian M, et al. Infrared and visible image fusion using saliency detection based on shearlet transform. In: *Proceedings of the 13th international computer conference on wavelet active media technology and information processing (ICCWAMTIP)*, Chengdu, China, 16–18 December 2016, pp.273–276. New York: IEEE.
5. Ma J, Zhou Z, Wang B, et al. Infrared and visible image fusion based on visual saliency map and weighted least square optimization. *Infrared Phys Techn* 2017; 82: 8–17.

6. Derlatka M and Bogdan M. Fusion of static and dynamic parameters at decision level in human gait recognition. In: *Proceedings of the international conference on pattern recognition and machine intelligence*, Warsaw, 30 June–3 July 2015, pp.515–524. Cham: Springer.
7. Al-Tayyan A, Assaleh K and Shanableh T. Decision-level fusion for single-view gait recognition with various carrying and clothing conditions. *Image Vision Comput* 2017; 61: 54–69.
8. Chiorean L and Vaida MF. Medical image fusion based on discrete wavelet transform using Java technology. In: *Proceedings of the ITI, international conference on information technology interfaces*, Dubrovnik, 22–25 June 2009, pp.55–60. New York: IEEE.
9. Wang N and Wang W. An image fusion method based on wavelet and dual-channel pulse coupled neural network. In: *Proceedings of the IEEE international conference on progress in informatics and computing (PIC)*, Nanjing, 18–20 December 2015, pp.270–274. New York: IEEE.
10. Kannan K and Perumal SA. Optimal decomposition level of discrete wavelet transform for pixel based fusion of multi-focused images. In: *Proceedings of the conference on computational intelligence and multimedia applications*, Sivakasi, India, 13–15 December 2007, vol. 3, pp.314–318. New York: IEEE.
11. Li M and Dong Y. Image fusion algorithm based on contrast pyramid and application. In: *Proceedings of international conference on mechatronic sciences, electric engineering and computer (MEC)*, Shenyang, China, 20–22 December 2013, pp.1342–1345. New York: IEEE.
12. He D, Meng Y and Wang C. Contrast pyramid based image fusion scheme for infrared image and visible image. In: *Proceedings of the geoscience and remote sensing symposium (IGARSS)*, Vancouver BC, Canada, 24–29 July 2011, pp.597–600. New York: IEEE.
13. Jin H, Liu F and Jiao LC. A method of image fusion based on multiscale contrast pyramid and directional filter banks. *Acta Electron Sinica* 2007; 35(7): 1295–1300.
14. Jin H, Jiao L, Liu F, et al. Fusion of infrared and visual images based on contrast pyramid directional filter banks using clonal selection optimizing. *Opt Eng* 2008; 47(2): 122–122.
15. O’rourke TP and Stevenson RL. Human visual system based wavelet decomposition for image compression. *J Vis Commun Image R* 1995; 6(2): 109–121.
16. Du J, Li W and Xiao B. Anatomical-functional image fusion by information of interest in local Laplacian filtering domain. *IEEE T Image Process* 2017; 26(12): 5855–5866.
17. Lan Z, Lin M, Li X, et al. Beyond Gaussian pyramid: multi-skip feature stacking for action recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Boston, MA, 7–12 June 2015, pp.204–212. New York: IEEE.
18. Olkkonen H and Pesola P. Gaussian pyramid wavelet transform for multiresolution analysis of images. *Graph Model Im Proc* 1996; 58(4): 394–398.
19. Bamberger RH and Smith MJT. A filter bank for the directional decomposition of images: theory and design. *IEEE T Signal Pr* 1992; 40(4): 882–893.
20. Do MN and Vetterli M. The contourlet transform: an efficient directional multiresolution image representation. *IEEE T Image Process* 2005; 14(12): 2091–2106.
21. Hof PR and Van Der Gucht E. Structure of the cerebral cortex of the humpback whale, *Megaptera novaeangliae* (Cetacea, Mysticeti, Balaenopteridae). *Anat Rec* 2007; 290(1): 1–31.
22. Mirjalili S and Lewis A. The whale optimization algorithm. *Adv Eng Softw* 2016; 95: 51–67.
23. Vaidyanathan PP. *Multirate systems and filter banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
24. Zhan K, Xie Y, Wang H, et al. Fast filtering image fusion. *J Electron Imag* 2017; 26(6): 063004.
25. Patil U and Mudengudi U. Image fusion using hierarchical PCA. In: *Proceedings of the international conference on image information processing*, Shimla, India, 3–5 November 2011, pp.1–6. New York: IEEE.