An Effective Image Fusion Method Based on Nonsubsampled Contourlet Transform and Pulse Coupled Neural Network

Lijuan Ma, Chunhui Zhao
College of Information and Communication Engineering
University of Harbin Engineering
Harbin, Heilongjiang Province, China
madanning.1988@163.com

***Abstract*—In order to solve the problem of spectral distortion and the fuzzy texture in visible and infrared image fusion technology, a novel visible and infrared image fusion method based on the Nonsubsampled Contourlet Transform (NSCT) and Pulse Coupled Neural Networks (PCNN) is proposed in this paper. First, we gain three components of visible image, luminance I, chrominance H and saturation S, using the IHS transform. Then, we gain three coefficients, low frequency sub-band, passband sub-band and high frequency coefficient by decomposing the component I and infrared image with the help of the NSCT. Next, we use weighted-sum method to fuse the low frequency sub-band and PCNN method to fuse the other sub-band coefficient respectively. At last, we gain the fusion image by using the inverse IHS transform on the fusion component I gained by the inverse NSCT transform. Experiments show that our method have better fusion quality and can be more better to keep the visible spectral and detail information than some traditional methods such as, Laplace method, Wavelet method and Lifting Wavelet method.**

Keywords — IHS transform; Nonsubsampled Contourlet transform; Pulse Coupled Neural Networks transform; image fusion

I. INTRODUCTION

Infrared image sensor is only sensitive to the thermal radiation of target scene, and the visible light image sensor is only sensitive to the reflection of target scene. We fused the visible image and infrared image in order to obtain more clearly, reliable and comprehensive image description of the same scene target. This method can effectively reduce the probability of false alarm and, improve the probability of correct identification and tracking of targets in complex background and interference conditions

It laid the foundation for accurate positioning and precise strikes of the target. Infrared and visible image fusion has important applications in aviation, remote sensing, national defense and many other areas.

Recent years, a lot of image fusion algorithms had been proposed, such as Laplace pyramid decomposition method, wavelet transform, lifting scheme of wavelets and contourlet transform which was developed in recent years. Contourlet transform as a multi-scale geometric analysis method solved the shortcoming of the wavelet transform which could not effectively represent the two-dimensional or higher singularity.

Recently, Cunha et al. proposed nonsubsampled contourlet transform (NSCT) [1-2]. This transformation has completely translational invariance so pseudo-Gibbs effect which once emerged in contourlet transform has already been solved.

Eckhom discovered stimulus-specific interactions between cell assemblies in cat primary visual cortex that could constitute a global linking principle for feature associations in sensory and motor systems [3], because of it's Global coupling and pulse synchronization, PCNN is widely used in image fusion [4].

In this paper, the author proposed an effective algorithm based on nonsubsampled contourlet and pulse coupled neural network for color visible and infrared image fusion. Simulation results certify that this fusion method is superior to that of Laplace, wavelet transform method and lifting scheme of wavelet transform method.

II. IHS TRANSFORM

Firstly, mapping the visible image to IHS space, common color space includes RGB color space and IHS color space [5]. RGB color space was commonly used for computer color monitor display system, wherein R represents red, G represents green and B represents blue. IHS color space contains three elements: the intensity, the hue and saturation. The relevance of these three elements is low, which enables us to separately processing them in the IHS space. IHS transform is defined as follow:

A. IHS transform

$$\begin{pmatrix} I \\ v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ -\sqrt{2}/6 & -\sqrt{2}/6 & 2\sqrt{2}/6 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \tag{1}$$

$$H = \arctan\frac{v_2}{v_1}, \quad S = \sqrt{v_1^2 + v_2^2} \tag{2}$$

Published by Atlantis Press, Paris, France.

B. *Inverse IHS transform*

$$\begin{pmatrix} R_{new} \\ G_{new} \\ B_{new} \end{pmatrix} = \begin{pmatrix} 1 & -1/\sqrt{2} & 1/\sqrt{2} \\ 1 & -1/\sqrt{2} & -1/\sqrt{2} \\ 1 & \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} I_{new} \\ v_1 \\ v_2 \end{pmatrix} \quad (3)$$

III. NSCT TRANSFORM

NSCT transform is divided into nonsubsampled pyramid filter bank (NSPFB) [6] and nonsubsampled directional filter bank (NSDFB) [7]. First using NSPFB decompose the image into different scales, and then applying NSDFB to decompose the sub-band image (except low frequency sub-band image) into different directions, thereby the sub-band image (coefficient) is obtained. The NSCT is a fully shift-invariant, multi-scale, and multi-direction expansion that has a fast implementation. The structure of NSCT is as follow:

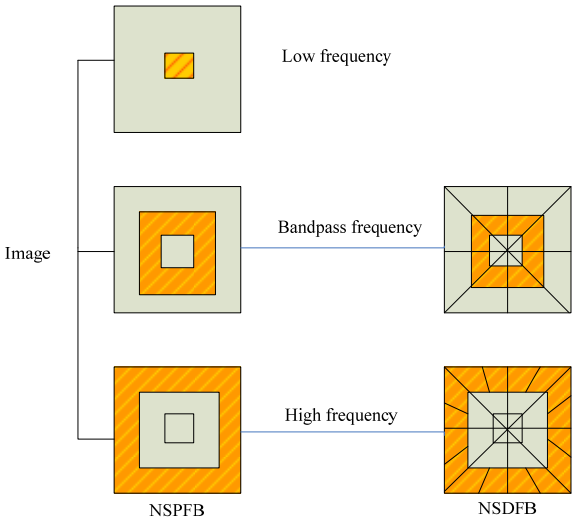

Figure 1. The structure of NSCT

A. *NSPFB*

The nonsubsampled Laplace pyramid which consists of two-channel nonsubsampled 2-D filter banks. Because the filter bank is not been nonsubsampled, thus it has the translational invariance. The ideal pass-band support of the low-pass filter at the $Y_A(m,n)$ th stage is the region $[-(\pi/2^j),\ (\pi/2^j)]^2$. Accordingly, the ideal support of the equivalent high-pass filter is the complement of the low-pass, i.e. the region $[(-\pi/2^{j-1}),(\pi/2^{j-1})]^2 / [(-\pi/2^j),(\pi/2^j)]^2$ [1].

B. *NSDFB*

Nonsubsampled directional filter banks are based on Bamberger and Smith constructed fan-shaped directional filter bank [8]. It is constructed by eliminating the downsamplers and upsamplers and then interpolation is done to the filter to get the same translational invariance nonsubsampled directional filter banks. Using nonsubsampled directional filter banks to decompose a certain scale sub-band image to n-level direction, it can obtain 2^n directional sub-band images which has same size with original image.

IV. PCNN TRANSFORM

PCNN is a feedback network which consists of a plurality of neurons connecting to each other, and each of neuron is composed of three parts: a receiving portion, a modulation portion and a pulse portion. For image processing, PCNN is a partially connected, monolayer and two-dimensional network. The number of neurons is equal to the number of input image pixels [9]. The structure of PCNN is shown as follow: Expression, as shown in Eq. (4).

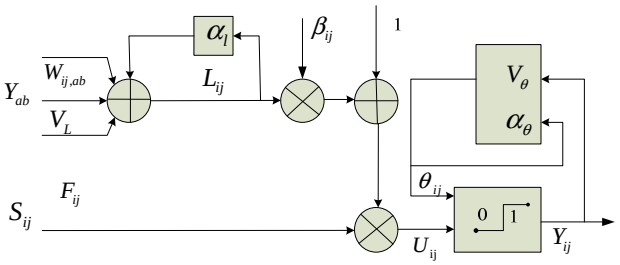

Figure 2. The structure of PCNN

Wherein $S(m,n)$ is the entry of the neurons and it is the pixel value at (m,n), $L(m,n)$ is the total linking input, $U(m,n)$ is the internal activity, $Y(m,n)$ is the output, $\theta(m,n)$ is threshold, $Y_A(m,n)$ is input time constant, α_θ is time decay constant, β is connection strength, w_{ij} is the weight coefficient matrix of the neurons, and k is the number of iteration. The external input signal $S(m,n)$ is directly used as the input of the neuron, V_θ is a pre-set threshold value and then decays exponentially. If $U_k(m,n) > \theta_k(m,n)$, neuron generates a pulse, it is called an ignition, after the k iteration. The number of ignition in (m,n) represents the information of this point.

V. FUSION RULE

In the process of image fusion, fusion rule is a key factor to obtain the high-quality fused images. In the following, we apply an effective fusion rule to image fusion.

A. *Image decomposition*

We get the luminance I of visible image by applying the IHS transform Eq. (1), define I component as *A* and infrared image as *B*. Then *A*, *B* are decomposed into three bands by

$$\begin{cases} F_k(m,n) = S_k(m,n) \\ L_k(m,n) = \exp(-\alpha_l) L_{k-1}(m,n) + V_l \sum_{a,b} W_{mn,ab} Y_{k-1}(m,n) \\ U_k(m,n) = F_k(m,n) \times (1 + \beta L_k(m,n)) \\ \theta_k(m,n) = \exp(-\alpha_\theta) \theta_{k-1}(m,n) + V_\theta Y_{k-1}(m,n) \\ Y_k(m,n) = step(U_k(m,n) - \theta_k(m,n)) \end{cases} \tag{4}$$

$$\begin{cases} C_{(i,j)}(m,n) = A_{(i,j)}(m,n) & |Y_A(m,n) - Y_B(m,n)| > \varepsilon \text{且} Y_A(m,n) > Y_B(m,n) \\ C_{(i,j)}(m,n) = B_{(i,j)}(m,n) & |Y_A(m,n) - Y_B(m,n)| > \varepsilon \text{且} Y_B(m,n) > Y_A(m,n) \\ C_{(i,j)}(m,n) = \left(A_{(i,j)}(m,n) + B_{(i,j)}(m,n)\right)/2 & |Y_A(m,n) - Y_B(m,n)| < \varepsilon \end{cases} \tag{7}$$

NSCT transform, low frequency sub-band is represented by $A(0)$, $B(0)$, passband sub-band and high frequency sub-band is represented by $A_{(ij)}$, $B_{(ij)}$. Wherein $A_{(ij)}$ is i th pass band, j th direction sub-band.

B. *For the low frequency*

Applying the algorithm below to fusion the low frequency sub-band

$$P_1 = \frac{V_1}{V_1 + V_2}, \quad P_2 = \frac{V_2}{V_1 + V_2} \tag{5}$$

$$C(0) = P_1 \times A(0) + P_2 \times B(0) \tag{6}$$

respectively, V_1, V_2 is the variance of $A(0)$, $B(0)$, $C(0)$ is the low frequency coefficients of fused image C.

C. *For the bandpass and high frequency*

Fusing bandpass frequency and high frequency sub-bands by using the method of the PCNN transform. After k iterations, obtained $Y_A(m,n)$ and $Y_B(m,n)$, fusion threshold is an arbitrarily small number. Fusion rules are shown as the Eq. (7).

D. *The inverse transform*

After the steps above, fused coefficients $C(0)$, $C_{(i,j)}(m,n)$ are generated. First of all, obtaining fused component I_{new} by the inverse NSCT transform. Then, the fused image C is gained by using the inverse IHS transform Eq. (3).

VI. SIMULATION

To evaluate the performance of the proposed image fusion method, two experiments have been performed on two sets of visible and infrared images; 'Octec' of size224*168, 'Trees' of size 800*600. For the proposed fusion rule, the experiment results show that a relatively better fused image can be obtained if $V_l = 0.2$, $V_\theta = 20$, α_l=1, $\alpha_\theta = 2$, $\varepsilon = 0.015$. Defined some parameter's initial values, $\theta_0(m,n) = 0$, $Y_0(m,n) = 0$, $L_0(m,n) = 0$, and W take on the values as follow:

$$W = \begin{bmatrix} 0.3536 & 0.4472 & 0.5 & 0.4472 & 0.3536 \\ 0.4472 & 0.7071 & 1 & 0.7071 & 0.4472 \\ 0.5 & 1 & 0 & 1 & 0.5 \\ 0.4472 & 0.7071 & 1 & 0.7071 & 0.4472 \\ 0.3536 & 0.4472 & 0.5 & 0.4472 & 0.3536 \end{bmatrix} \tag{8}$$

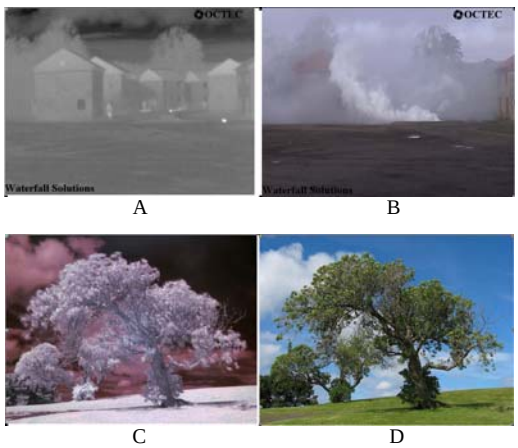

Figure 3. (A), (B) Octec. (C), (D) Trees. On the left is the infrared image, on the right is the visible image

The performance of the proposed method is compared with the conventional IHS-Laplace method, IHS-DWT method in [5] and IHS-lifting DWT method in [10].

TABLE I PERFORMANCE COMPARISON ON OCTEC

Algorithms	IHS-laplace	IHS-DWT	IHS-lifting DWT	Our algorithm
Entropy	7.5776	7.5743	7.5773	7.5815
Average correlation coefficient	0.0951	0.1052	0.1051	0.1055
SF	19.6007	18.8234	19.5137	19.8163

TABLE II PERFORMANCE COMPARISON ON TREES

Algorithms	IHS-laplace	IHS-DWT	IHS-lifting DWT	Our algorithm
Entropy	7.2430	7.2232	7.1863	7.2917
Average correlation coefficient	0.1051	0.1237	0.1233	0.1263
SF	24.4762	23.8466	24.4205	24.5123

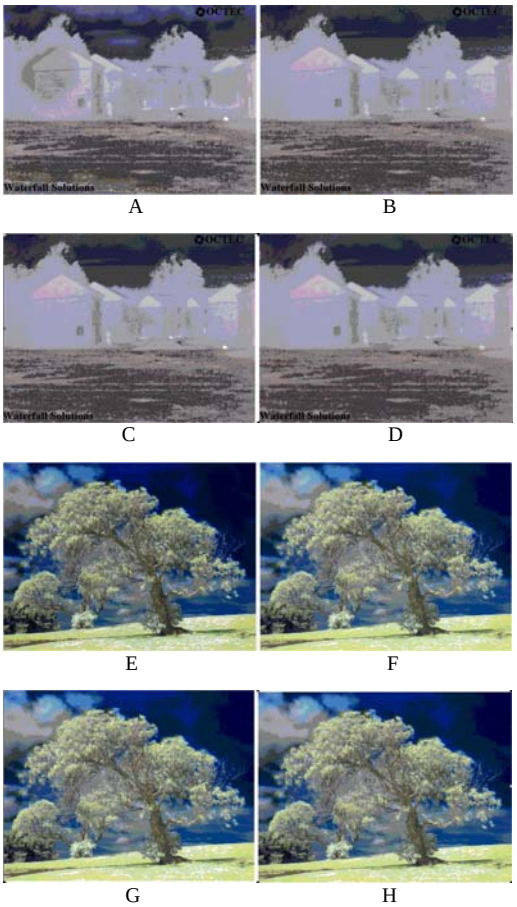

Figure 4. (A), (E) the results of IHS-laplace. (B), (F) the results of IHS-DWT. (C), (G) the results of IHS-lifting DWT. (D), (H) the results of proposed algorithm in this paper.

We can conclude from Fig.4.Picture (A) and (E) that they have obvious spectral distortion, in picture (B), (F), (C) and (G), the details blur, and picture (D), (H) with better spectral performance and clearer details.

VII. EVALUATION

Currently, there is no unified objective criterion to measure the quality of the fused images. In this paper, we use the following objective criteria to evaluate the quality of the fused images.

A. *Entropy*

Information entropy is an important indexes for the evaluation of images, and is defined by

$$H = -\sum_{i=0}^{L-1} P(i)\log_2 P(i) \tag{9}$$

in which $P(i)$ is the probability of the pixel value i. Higher entropy indicates more information in the fused image.

B. *Average correlation coefficient*

It reflects the similarity and spectral maintain performance of the fused image with the two original images. Higher average correlation coefficient indicates more information in the fused image.

$$C = \left(C_{(AC)} + C_{(BC)}\right)/2 \tag{10}$$

$$C_{(XY)} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij}-\overline{X}\right)\left(Y_{ij}-\overline{Y}\right)}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij}-\overline{X}\right)^2}\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(Y_{ij}-\overline{Y}\right)^2}} \tag{11}$$

C. *Spatial Frequency*

Spatial frequency is an evaluation of image clarity variable, reflecting the activity level in the image. The definition is:

$$SF = \sqrt{\frac{1}{MN}\left[\sum_{i=1}^{M-1}\sum_{j=1}^{N-1}\left[F(i,j)-F(i,j-1)\right]^2+\left[F(i,j)-F(i-1,j)\right]^2\right]} \qquad (12)$$

Figures shown in table.1 and table2 indicate that the method proposed in this paper makes fused image obtained maximum entropy，highest average correlation coefficient and spatial frequency.

VIII. CONCLUSIONS

From the above analysis and experimental results, we can draw a conclusion that the proposed method takes the advantages of more valuable information and details of the outstanding characteristics contained in fused images, as well as better visual effects. However, this method takes long time, and not suitable for real-time processing. In order to balance fusion time and fusion effect, the high frequency was decomposed to four directions. The method just has one more direction compared with DWT and lifting-DWT method, so the fusion effect promotion is not particularly large. NSCT transform can decomposed image into any amount direction, the more direction of the division are made, the better fusion effect will be got, and combined with the integration of PCNN will need more time.

ACKNOWLEDGMENT

This study was partially supported by the National Natural Science Foundation of China(Grant No. 61077079), and by the Ph.D. Programs Foundation of Ministry of Education of China(Grant No. 20102304110013)，and by the key program of Heilongjiang Nature Science foundation, and by the International Exchange Program of Harbin Engineering University for Innovation-oriented Talents Cultivation. .

REFERENCES

[1] L. D. Cunha, J. P. Zhou and N. D. Minh, “The nonsubsampled contourlet transform: Theory, design, and applications,” IEEE Transl. Image Processing, vol. 15, no. 10, pp. 3089–3101, October 2006.

[2] R. Eslami and H. Radha, “Translation-invariant contourelt transform and its application to image denoising,” IEEE Transl. Image Processing, vol. 15, no. 11, pp. 3362–3374, November 2006.

[3] R. Eckhom, H. J. Reitboeck, M. Arndt and P. Dicke, “Feature linking via synchronization among distributed assemblies,” Mit Press Journals, vol. 2, no. 3, pp. 293–307, 1990.

[4] B. C. Xu and Z. Chen, “A multi-sensor image fusion algorithm based on PCNN,” WCICA2004, pp. 3679–3682, June 2004.

[5] C. X. Lu, Q. Pan, Y. M. Cheng, “New image fusion method based on HIS and wavelet transform,” Application Research of Computers, vol. 25, no. 2, pp. 3690–3695, February 2008.

[6] L. Tang, F. Zhao and Z. G. Zhao, “The nonsubsampled contourlet trams-form for image fusion,” ICWAPR07, pp. 305–310, November 2007.

[7] G. F. Xie, J. W. Yan, Z. Q. Zhu and B. G. Chen, “Image fusion algorithm based on neighbors and cousins information in nonsubsampled contourlet tramsform,” ICWAPR07, pp. 1797–1802, November 2007.

[8] R. H. Bamberger, “A filter bank for the directional de-composition of images:Theory and design,” IEEE Transl. Signal Processing, vol. 40, no. 4, pp. 882–893, 1992.

[9] J. L. Johnson, D. Ritter. “Observation of periodic waves in a pulse-coupled neural network,” Optics Lett, vol. 18, no. 15, pp. 1253–1255, 1993.

[10] H. Chen and Y. Y. Liu, “An infrared image fusion algorithm based on lifting wavelet transform,” Laser and Infrared, vol. 39, no. 1, January 2009.