Fusion of Infrared and Visible Images based on NSCT and Modified PCNN

Xue-yan Zhou*, Jia-min Gong and Ren-ping Xing

School of Electronics Engineering, Xi’an University of Posts and Telecommunications, 710121 Xi’an, China

ABSTRACT

For purpose of improving the quality of multi-source fusion images, this paper introduces a fusion method of infrared and visible images based on the combination of the non-subsampled contourlet transform (NSCT) and the improved pulse coupled neural network model (PCNN). To begin with, registered images are decomposed by NSCT, the low frequency and high frequency sub-bands are also obtained. Then, the energy of Laplacian (SML) is used for the link intensity of PCNN, computing spatial frequency (SF) with a low frequency sub-band, and the gradient energy (EOG) with a high frequency sub-band. By means of treating them as external input of PCNN. Then fused coefficients are selected by sum of ignition output amplitude maximum, a fused image is finally obtained by inverse NSCT. Experimental results show that the fusion image has a high definition and effectively preserves information of the source image.

INTRODUCTION

Image fusion can be divided into three levels, that is, pixel-level fusion, feature-level fusion and decision-level fusion. The pixel-level fusion is conducted on the directly collected initial information, which is in favor of obtaining the richer and more accurate details of the scene [1,2]. The commonly-seen scale decomposition methods are based on the analysis of multiple scales and directions. The common decomposition methods are made up of wavelet transform and pyramid decomposition; while an expansion from one-dimensional to two-dimensional wavelet transform can merely capture the horizontal, diagonal, as well as vertical-three directions of the information, and it cannot stand for the information of high frequency in all directions and scales[3,4]. In 2002, Do has proposed the Contourlet transform, which is directly discrete data-oriented. It possesses the good local characteristics of time-frequency and multiple scales, but in the meanwhile, it can realize the decomposition in the high frequency parts from arbitrary directions, and provide edge information from arbitrary directions. Nevertheless, during the two decomposition processes, there is a need to engage in sub-sampling operation. Therefore, the Gibbs effect is come into being, bringing about the distortion of the fusion results. In 2006, Non-Subsampled
Contourlet transform (NSCT) has been proposed on the basis of the Contourlet transform, which has not only enhanced the analysis of Contourlet in multi-directivity, but also overcome the distortion of fusion that is generated from the Gibbs effect, and thus, a high recognition has been gained [5,6].

Pulse coupled neural networks (PCNN) was developed by Eckhorn et al. It is a model that simulates the information processing of neuronal cells in the visual area of mammals. Owing to its single-layer neural network model, aspects such as image segmentation and edge detection can be achieved without learning or training. In recent years, some scholars have combined NSCT with PCNN for image fusion, and thus, good experimental results have been obtained[7].

**NSCT**

Contourlet is a two-layered and indivisible direction filter bank structure. In the discrete domain, firstly, Laplacian pyramid (LP) is adopted along the smooth edge of the target object so as to capture the different directions of the isolated singular points. Then it links the isolated singular points in the same direction to a basic contour segment by means of the direction filter group (DFB), it can effectively and flexibly present the detailed information of images in any direction. On account of Contourlet has caused the Gibbs effect and then brought about the distortion of the fusion result, the proposal of NSCT seems to be of necessity. NSCT is a translation-invariant Contourlet transform that makes up of a non-subsampled Laplacian pyramid (NSP) and a non-subsampled directional filter banks (NSDFB). NSCT transform is equipped with good time-frequency analysis, and at the same time, it is an efficient tool for image processing in the two-dimensional discrete domain. The structure of NSCT is demonstrated in Figure 1[8].

![Figure 1. Decomposition framework of NSCT.](image)

**PCNN**

The neuronal activity in the pulse coupled neural network (PCNN) is the approximation of the human visual neurons. The input of the traditional neurons in the field is the algebraic sum of the weighted inputs, but the input excitation of the PCNN
model is the combination of both the external input image signal and the connection input, which makes PCNN show its unique superiority in the image processing, as well as guarantee the integrity of the information [9]. The structure of PCNN is shown in Figure 2.

As the PCNN model involves a large number of parameters, it is very complicated to list all of them. For the sake of making the image fusion more efficient, this paper uses a simplified PCNN model that has proposed in the literature [9], and its mathematical expression is demonstrated as follows:

\[
\begin{align*}
F_{ij} (n) &= S_{ij} \\
L_{ij} (n) &= \exp(-\alpha_L) L_{ij} (n-1) + V_L \sum_{pq} W_{ijpq} Y_{pq} \\
U_{ij} (n) &= F_{ij} (n) (1 + \beta L_{ij} (n)) \\
\theta_{ij} (n) &= \exp(-\alpha_\theta) \theta_{ij} (n-1) + V_\theta Y_{ij} (n) \\
Y_{ij} (n) &= \begin{cases} 
1 & U_{ij} (n) \geq \theta_{ij} \\
0 & U_{ij} (n) < \theta_{ij}
\end{cases} \\
T_{ij}^{lk} &= T_{ij}^{lk} (n-1) + Y_{ij}^{lk} (n)
\end{align*}
\]

(i, j) refers to the grey level (the label of the neuron) of the corresponding pixel, F_{ij} is the feedback input, n is the number of iterations, as well as S_{ij} is the external excitation of the neuron’s forced excitation. L stands for the link input of the neuron, V_L and V_F respectively represent the amplification coefficient of feedback signals from input signals, and the weight matrix W is the matrix of the connection weights between the neurons; k and l represent the connection range of the neuron with its adjacent neurons. U_{ij} is the internal activity of the neuron; \beta is the connection intensity coefficient, which determines the contributions of linear link input to the internal activities; \theta_{ij} (n) is the output of the variable threshold function; V_\theta and \alpha_\theta are the amplification coefficients of the variable threshold function signal and the decay time constant respectively.
THE METHOD OF THIS PAPER
Determination of External Input Excitation in PCNN Model

In the traditional PCNN-based image fusion, in accordance with all the experiments, the connection intensity coefficient $\beta$ of each neuron is the same fixed value. Nevertheless, according to the sensitivity of the visual system of the human eye towards the edge, the connection intensity of each neuron is impossible to be exactly the same [10]. Therefore, this paper selects the SML of the sub-band coefficient as the adaptive connection intensity of PCNN [12]. The low-frequency component is the approximation of the source image, and thus, the quality of the fusion result will be obviously affected by the low-frequency fusion algorithm. In this paper, the author uses the SF of the low sub-band coefficient as the external input of PCNN so as to obtain the low frequency fusion coefficient. SF reflects the degree of change in the spatial point of each pixel. The larger the value SF is, the clearer the image will be [11].

The high frequency part is mainly adopted to characterize the edge, texture, as well as other details. This paper suggests EOG as the external input of PCNN so as to get a high-frequency fusion coefficient, in which the formula is listed as follows:

$$ML(i, j) = |2D(i, j) - D(i - 1, j) - D(i + 1, j)| + |2D(i, j) - D(i, j - 1) - D(i, j + 1)|$$  \hspace{1cm} (2)

$$SML(i, j) = \sum_{p=-P}^{P} \sum_{q=-Q}^{Q} ML^2(i+p, j+q)$$  \hspace{1cm} (3)

$$SF(i, j) = \left( \frac{1}{(2P+1)(2Q+1)} \sum_{m=-P}^{P} \sum_{n=-Q}^{Q} (D(m,n+1)-D(m,n))^2 + (D(m+1,n)-D(m,n))^2 \right)^{\frac{1}{2}}$$  \hspace{1cm} (4)

$$G(i, j) = \left( (D_{l,k}(i, j) - D_{l,k}(i+1, j))^2 + (D_{l,k}(i, j) - D_{l,k}(i, j+1))^2 \right)^{\frac{1}{2}}$$  \hspace{1cm} (5)

$$EOG_{l,k}(i, j) = \sum_{i=p, j=q} |G_{l,k}(i, j)|$$  \hspace{1cm} (6)

$$D_{l,k}^{T}(i, j) = \begin{cases} T_A(i, j) & T_A(i, j) > T_B(i, j) \\ T_B(i, j) & T_A(i, j) \leq T_B(i, j) \end{cases}$$  \hspace{1cm} (7)

$D_{l,0}(i, j)$ refers to the low frequency coefficient value at $(i, j)$; $D_{l,k}(i, j)$ represents the high frequency coefficient value at $(i, j)$; $G_{l,k}(i,j)$ is the gradient of the sub-band coefficient; $ML(i,j)$ is Laplacian; $l$ and $k$ are numbers of the decomposed sub-bands and the directions of sub-bands respectively; $(2P + 1) \times (2Q + 1)$ is the size of sliding windows, $3 \times 3$ is the window selected by the text; $T_A(i, j)$ and $T_B(i, j)$ are the ignition matrix that correspond to each sub-band obtained by the ignition process.

93
Fusion Process and Algorithm Procedure Description

Firstly, the low frequency sub-band coefficient and the high frequency sub-band coefficient are obtained by the NSCT transformation. Secondly, the author respectively calculates the SF of the low frequency sub-bands and the EOG of the high frequency sub-bands as the external input of PCNN model. Finally, the author gets the ignition matrix of the source image, and as well, the fusion coefficient is determined by the number of times of ignition. The final fusion image is transformed into the fusion image by NSCT inversion. The fusion process is displayed in Figure 3:

![Figure 3. The framework of image fusion in this paper.](image)

SIMULATION AND EXPERIMENTAL RESULTS ANALYSIS

In order to verify the authenticity of the algorithm, two sets of visible and infrared images is used to test and to compare with three traditional classical algorithms. Method 1 is Contourlet algorithm. Method 2 is based on the simple NSCT, that is, the low-frequency part adopts the average method and the high-frequency part uses the larger absolute value. Method 3 is a combination of another NSCT and PCNN model, in which the low-frequency part uses the region energy maximum; while the high-frequency part treats SML as an external input excitation to PCNN. Method 4 is the algorithm that has proposed in this paper.

In this paper, all the experimental methods that NSCT transformation parameter setting are: NSP adopts "9-7" biorthogonal, NSDFB adopts "pkva" ladder filter, and the number of decomposition layers is 4 layers, the number of directions is 2,2,4,4.

PCNN parameter settings: $\alpha_L = 1, \alpha_o = 0.8, V_L = 1, V_o = 20, \beta = 3, n_{\text{max}} = 100$. 
Experimental results are demonstrated in Figure 4. Generally speaking, the visual effects of Figure (c) are relatively vague, even if the background and character information can be easily identified, the differences between the gray degrees are not so obvious and the overall level is poor. The visual effects of Figure(d) have been improved, but unfortunately, the edge of the image is blurred. This fusion method based on a single pixel value ignores the correlation between the pixels, and thus, leads to a reduction in the contrast of the target edge in the fused images. Figure(e) can better keep the contours around the person, but the visual effect of the leaf area below the graph is still blurred. With regard to the experiment method in this paper, on the one hand, the background and target characters are well identified; on the other hand, the edge details of the leaf area in the lower left corner of the source image are well preserved.

![Image](image_url)

(a) Visible image  (b) Infrared image  (c) Contourlet
(d) NSCT-PCNN  (e) NSCT-SML-PCNN  (f) The method of this paper

Figure 4. The results of the fusion of different algorithms.

The second experimental results are shown in Figure 5. The visual effects of Figure (d) and Figure (e) are relatively vague and the contrast of the fused image is reduce. The visual effects of Figure (d) and Figure (e) is well and retain many detailed edge feature of the original image.

![Image](image_url)

(a)Visible image  (b)Infrared image  (c)Contourlet
(d) NSCT-PCNN  (e) NSCT-SML-PCNN  (f) The method of this paper

Figure 5. The results of the fusion of different algorithms.
In order to make the evaluation of the fusion image more accurate, subjective evaluation and objective evaluation should be combined. This paper compares the fusion images by adopting the entropy (E), mutual information (MI), average gradient (AG), as well as Standard Deviation (SD). Entropy is used to measure the abundance of the image information. With the increase of the entropy, the more obvious the effect of fusion is, and the more abundant information is contained from the source image. The mutual information manifests the sum of the information that obtained from the source image. The greater the mutual information is, the greater correlation is between the fusion image and the source image. Besides, the average gradient describes the clarity of the image, and the greater the average gradient is, the more prominent the detailed information is. Moreover, the standard deviation depicts the degree of dispersion between the image pixel gray scale, and the larger the average gray scale is, the better fusion effect is [12]. It can be seen from TABLE I that the method of this paper is better than the other three methods. Even if the mutual information declines, the other evaluation indexes are better than those of the other three methods. At the same time, the fusion image clarity has been improved and further validates the effectiveness of the proposed method.

<table>
<thead>
<tr>
<th>Test 1</th>
<th>Method</th>
<th>Evaluation index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>Contourlet</td>
<td>6.7567</td>
</tr>
<tr>
<td></td>
<td>NSCT-PCNN</td>
<td>6.8630</td>
</tr>
<tr>
<td></td>
<td>NSCT-SML-PCNN</td>
<td>6.9415</td>
</tr>
<tr>
<td></td>
<td>The method of this paper</td>
<td>7.0117</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test 2</th>
<th>Method</th>
<th>Evaluation index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>Contourlet</td>
<td>7.2551</td>
</tr>
<tr>
<td></td>
<td>NSCT-PCNN</td>
<td>6.1441</td>
</tr>
<tr>
<td></td>
<td>NSCT-SML-PCNN</td>
<td>6.9415</td>
</tr>
<tr>
<td></td>
<td>The method of this paper</td>
<td>7.3256</td>
</tr>
</tbody>
</table>

**CONCLUSIONS**

Based on the characteristics of NSCT and PCNN, this paper presents an image fusion method that combines NSCT and the improved PCNN. Through the comparison of several experimental results, it can be concluded that the proposed algorithm preserves the edge feature of the source image as much as possible, achieving the purpose of improving the quality of fusion results. Nevertheless, there are still some shortcomings in the method of this paper, for instance, the mutual information is declined compared to other methods. Therefore, further study is still required.
REFERENCES


