Remote Sensing Image Fusion via Boundary Measured Dual-Channel PCNN in Multi-Scale Morphological Gradient Domain

Tan, Wei; Xiang, Pei; Zhang, Jiajia; Zhou, Huixin; Qin, Hanlin

Published in:
IEEE Access

DOI:
10.1109/ACCESS.2020.2977299

Publication date:
2020

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Remote Sensing Image Fusion via Boundary Measured Dual-Channel PCNN in Multi-Scale Morphological Gradient Domain

WEI TAN\textsuperscript{1,2}, PEI XIANG\textsuperscript{1}, JIAJIA ZHANG\textsuperscript{1}, HUIXIN ZHOU\textsuperscript{1}, AND HANLIN QIN\textsuperscript{1}

\textsuperscript{1}School of Physics and Optoelectronic Engineering, Xidian University, Xi’an 710071, China
\textsuperscript{2}Department of Computer Science, University of Copenhagen, 2100 Copenhagen, Denmark

Corresponding author: Wei Tan (twtanwei1992@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61675160, in part by the 111 Project under Grant B17035, and in part by the China Scholarship Council under Grant CSC201906060047.

\section*{ABSTRACT}
In this paper, a remote sensing image fusion method based on boundary measured dual-channel pulse-coupled neural network (PCNN) in multi-scale morphological gradient (MSMG) domain is proposed. Firstly, the panchromatic (PAN) image is decomposed into three parts, a small-scale image, a large-scale image, and a base image through a co-occurrence filtering (CoF)-based decomposition model. Secondly, an HSI transform is applied in the multispectral (MS) image to obtain intensity, hue and saturation components. Thirdly, a PCNN fusion strategy modulated by MSMG is used to fuse the base image and the intensity component of the MS image. Then, a fused intensity image is obtained by combining the small-scale image, large-scale image and the fused approximate image. Finally, the final fused image can be reconstructed by an inverse HSI transform. Experiments in four datasets demonstrate that the proposed method obtains the best performance in most cases.

\section*{INDEX TERMS}
Remote sensing image fusion, co-occurrence filtering, multi-scale morphological gradient, pulse-coupled neural network.

\section*{I. INTRODUCTION}
Nowadays, with the rapid development of remote sensing technology, there are more types of remote sensing images available when imaging a view [1], [2]. As a widely used remote sensing image, multispectral (MS) images containing plentiful spectral information are applied in the fields of disaster monitoring [3], marine research [4], object recognition [5], etc. However, the disadvantage of low spatial resolution makes it look blurry. The characteristic of a panchromatic (PAN) image is just the opposite of an MS image, which has high spatial resolution and low spectral resolution. Remote sensing images obtained from a single sensor have the disadvantage of incomplete information. Therefore, the MS and PAN image fusion (also called pansharpening) technology is crucial for obtaining an integrated image which contains plenty of spatial and spectral information.

In the past few decades, MS and PAN image fusion technology developed very rapidly. This technology is mainly divided into two categories: component substitution (CS) based algorithms [6]–[9] and multi-scale decomposition (MSD) based algorithms [10]–[14]. For the CS based algorithm, the MS images are decomposed into different transform domain, such as color/intensity space or principal components. Then the color/intensity spaces or first principal components are substituted by the corresponding contents of the PAN image. Finally, the fused image is reconstructed through the corresponding inverse transform. Classical CS based algorithms contain hue saturation intensity (HSI) transform [6], principal component analysis (PCA) [7], Gram-Schmidt (GS) [9], etc. The framework of these algorithms is very simple and can achieve efficient fusion performance in spatial resolution. However, since the bandwidth of PAN image cannot be completely covered by the MS images, some spectral distortions may occur.

Compared with the CS based algorithms, MSD based algorithms developed very rapidly. Famous MSD algorithms include: wavelet transform [10], curvelet transform [12], contourlet transform [13] and shearlet transform (NSST) [14]. In this kind of algorithms, the source images are first...
decomposed into different scales of sub-images through corresponding MSD methods. Then the sub-images are merged based on different fusion rules in different sub-bands. Finally, an inverse MSD transform is used to reconstruct the fused image. The MSD based algorithms can achieve better performance in spectral domain. However, the spatial resolution of MSD based algorithms is lower than that of the CS based algorithms.

In recent years, edge-preserving filtering (EPF) algorithms have achieved good progress in image fusion field. The EPF has the characteristics of maintaining the edge structure. Meanwhile, some fine edges will be smoothed off [15]. According to this feature, a novel hybrid decomposition model is produced [16]–[18]. An EPF and Gaussian filtering (GF) is used in this model. Using the edge-preserving characteristic of EPF and the smoothing characteristic of GF, the input image is decomposed into three layers: small-scale layer, large-scale layer, and base layer. Classical EPF-GF decomposition methods include Bilateral-Gaussian filtering decomposition (BGFD) [16], Rolling guidance-Gaussian filtering decomposition (CGGFD) [17], and Curvature-Gaussian filtering decomposition (CGFD) [18]. This hybrid decomposition model has obtained great success in image fusion. Therefore, this type of method is an important branch of future image fusion research.

Inspired by the advantages of EPF, Jevnisek and Avidan proposed a novel EPF called co-occurrence filtering (CoF) [19]. Its biggest advantage is that the sharp details in a local region can be preserved when smoothing the fine textures. Since this characteristic, the edges within a local region can be preserved while the image is smoothed.

Over the past few years, neural networks have been widely used in many fields [20], [21]. As a method of stimulating the visual principle structure of cats, pulse-coupled neural network (PCNN) is a simplified neural network model. However, since one pixel corresponds to one neuron in the PCNN model, human eyes are not sensitive to edge information when using PCNN alone. Therefore, it is necessary to utilize a feature operation before the PCNN processing. The most commonly used feature operation is spatial frequency (SF) [22]. However, local block effect and blurring exist in the fused images, for the reason of inappropriate weighted coefficients of the sub-bands fusion.

In order to overcome the disadvantages, an effective remote sensing image fusion via CoF and multi-scale morphological gradient domain dual-channel PCNN is proposed in this paper. Firstly, the PAN image is decomposed into small-scale image, large-scale image and base image through a CoF and a GF. Secondly, an HSI transform is applied in the MS image to obtain intensity, hue and saturation components. Thirdly, the base image and the intensity component of MS image are merged in the light of multi-scale morphological gradient (MSMG) dual-channel PCNN fusion rule to obtain a fused approximate image. Then, combining the small-scale image, large-scale image and the fused approximate image to obtain a fused intensity image. Finally, the fused image can be reconstructed by an inverse HSI transform. Experimental results demonstrate that the proposed algorithm outperforms many state-of-the-art algorithms.

The main contributions of this proposed algorithm are as follows:

1. The effective edge-preserving method CoF is first employed in the field of remote sensing image fusion. And the CoF-based image decomposition model is also presented in this paper.

2. An effective sub-image fusion rule based on MSMG dual-channel PCNN (MSMG-PCNN) model is employed in this paper. The gradient information of the image can be easily extracted, and the size of structure can be easily changed to adapt to the scale of structure.

3. The proposed CoF-MSMG-PCNN achieves the best performance in most cases in qualitative and quantitative.

The rest of this paper is organized as follows. The proposed algorithm which contains the theory of CoF and MSMG-PCNN is described in Section II. The experimental results are given in Section III, and the parameter discussion is given in Section IV. Finally, the conclusion is drawn in Section V.

II. PROPOSED METHOD

Remote sensing image fusion aims to produce an integrated image which contains plenty of spatial and spectral information. To realize the target, an effective remote sensing image fusion algorithm through CoF and MSMG-PCNN (we call it CoF-MSMG-PCNN) is proposed.

![FIGURE 1. Framework of the CoF-MSMG-PCNN fusion method.](image-url)

The framework of the proposed algorithm is shown in Fig. 1. It can be seen that the PAN image is decomposed...
into small-scale image (S), large-scale image (L) and base image (B). Then the intensity image (I) of MS image and B is merged into a fused approximate image (FA) through MSMG-PCNN fusion rule. The fused intensity image (FI) is reconstructed by S, L and FA. Finally, the fused image is obtained by an inverse HSI transform.

In this section, the proposed algorithm, along with the required CoF and MSMG-PCNN theory, will be introduced.

A. CO-OCCURRENCE FILTERING

1) THEORY OF CoF

Co-occurrence filtering is an effective EPF which is proposed by Jevnisek and Avidan. Similar with the BF, pixel values which co-occur frequently will be smoothed, while the pixel values co-occur rarely will be preserved. The CoF is defined as

\[ V_p = \frac{\sum_{q \in N(p)} w(p, q) \cdot U_q}{\sum_{q \in N(p)} w(p, q)} \]  

where \( V_p \) and \( U_q \) are output and input, \( p \) and \( q \) are pixel indices. \( w(p, q) \) is the weigh between input pixel \( p \) and output pixel \( q \).

In GF, \( w(p, q) \) is formed as

\[ w(p, q) = G_{\sigma_s}(p, q) \cdot \exp(-\frac{d(p, q)^2}{2\sigma_s^2}) \]  

where \( G_{\sigma_s}(p, q) \) denotes GF, \( d(p, q) \) is the Euclidean distance between pixels \( p \) and \( q \). And \( \sigma_s \) is the parameter which is related with the distribution of Gaussian function. In BF, \( w(p, q) \) is defined as

\[ w(p, q) = G_{\sigma_s}(p, q) \cdot \exp(-\frac{|I_p - I_q|^2}{2\sigma_r^2}) \]  

where \( \sigma_r \) is a user designed coefficient.

In CoF, the GF is replaced with a normalized co-occurrence matrix. It takes the form of

\[ w(p, q) = G_{\sigma_s}(p, q) \cdot M(I_p, I_q) \]  

where \( M \) is a 256 × 256 matrix, which is computed by

\[ M(m, n) = \frac{C(m, n)}{h(m)h(n)} \]  

where \( C(m, n) \) is obtained through a co-occurrence matrix \( C(m, n) \). The values \( m \) and \( n \) denote the grayscale level, and \( h(m) \) and \( h(n) \) represent the corresponding frequency. \( C \) is defined as

\[ C(m, n) = \sum_{p,q} \exp(-\frac{d(p, q)^2}{2\sigma^2}) |I_p = m| |I_q = n| \]  

where \( \sigma \) is a designed coefficient, and the operation \( [] \) means if the expression inside is true, the value will equal 1, otherwise the value equals 0.

Fig. 2 shows two examples of CoF. It can be seen that grass in (a) and the background in (b) are smoothed and removed, while the strong structures and edges are preserved.

2) CoF-BASED DECOMPOSITION METHOD

Since CoF preserves edges effectively, it can be used in image decomposition. In this paper, an image decomposition model based on CoF and GF is presented. An input image can be decomposed into three sub-images in different scales, through this decomposition method. As shown in Fig. 3, the process is divided into three steps: Firstly, the CoF and GF are used in the input image \( J \) to obtain filtered images \( I_c \) and \( I_g \), respectively. Secondly, the small-scale image (S) is obtained by the difference between \( J \) and \( I_c \), and the large-scale image (L) is also the differential result.
between $J$ and $I_g$. Finally, the base image (B) is represented by $I_g$. The decomposition process can be formulated as

$$I_g = GF(J, \sigma, \rho)$$  \hspace{1cm} (7)

$$I_c = CoF(J, p)$$  \hspace{1cm} (8)

$$S = J - I_c$$  \hspace{1cm} (9)

$$L = I_c - I_g$$  \hspace{1cm} (10)

$$B = I_g$$  \hspace{1cm} (11)

where $GF(\cdot)$ and $CoF(\cdot)$ denote the Gaussian filtering and co-occurrence filtering operations, respectively. $\rho$ and $\delta$ represent the radius and variance of the GF. In this paper, if both the parameters are large, the fused image will appear unexpected block, and if they are small, the fused image will not be enough clear. Therefore, these two parameters are set to $\rho = 4$, $\delta = 10$. $p$ represents the size of filter window. Similar with $\rho$ and $\sigma$, $p$ is set to 10.

**B. MULTI-SCALE MORPHOLOGICAL GRADIENT DUAL-CHANNEL PCNN**

1) MSMG OPERATOR

Multi-scale morphological gradient is an effective operator which can easily and well extract gradient information from an image. Since its superior characteristics, MSMG has been employed as a type of focus measure in multi-focus image fusion [23]. The specific details are described below.

Firstly, the multi-scale structuring element is defined as

$$SE_t = SE_1 \oplus SE_1 \oplus \cdots \oplus SE_1, \quad t \in \{1, 2, \ldots, N\}$$  \hspace{1cm} (12)

where $SE_1$ denotes a basic structure element, and $t$ represents the number of scales. In this paper, the size of the structure element is set to 13, and $t$ is set to 7.

Secondly, the gradient feature $G_t$ can be represented by the morphological gradient operators from the image $f$

$$G_t(x, y) = f(x, y) \ominus SE_t - f(x, y) \oplus SE_t$$  \hspace{1cm} (13)

where $\ominus$ and $\oplus$ denote the morphological dilation and erosion operators, respectively. $(x, y)$ denotes the pixel coordinate.

Thirdly, MSMG is obtained by computing the weighted sum of gradients in all scales.

$$M(x, y) = \sum_{t=1}^{N} w_t \cdot G_t(x, y)$$  \hspace{1cm} (14)

where $w_t$ represents the weight of gradient in $t$-th scale, and it can be represented as

$$w_t = \frac{1}{2t + 1}$$  \hspace{1cm} (15)

Fig. 4 shows the results of MSMG. It can be seen that the gradient information of the images have been well extracted.

2) CLASSICAL PCNN MODEL

As the third-generation artificial neural network, PCNN has achieved great success in image fusion field. A PCNN model often contains three parts: the receptive field, the modulation field and the pulse generator. The expressions of a simplified dual-channel PCNN model can be defined as

$$F_{ij}^1(k) = S_{ij}^1(k)$$  \hspace{1cm} (16)

$$F_{ij}^2(k) = S_{ij}^2(k)$$  \hspace{1cm} (17)

$$L_{ij}(k) = \begin{cases} 1, & \text{if } \sum_{r, t \in S} Y_{rt}(k - 1) > 0; \\ 0, & \text{otherwise}. \end{cases}$$  \hspace{1cm} (18)

$$U_{ij}(k) = \max\{F_{ij}^1(k)(1 + \beta_{ij}^1 L_{ij}(k)), F_{ij}^2(k)(1 + \beta_{ij}^2 L_{ij}(k))\}$$  \hspace{1cm} (19)

$$Y_{ij}(k) = \begin{cases} 1, & \text{if } U_{ij}(k) \geq \theta_{ij}(k - 1); \\ 0, & \text{otherwise}. \end{cases}$$  \hspace{1cm} (20)

$$\theta_{ij}(k) = \theta_{ij}(k - 1) - d_e + V_{\theta} Y_{ij}(k)$$  \hspace{1cm} (21)

$$T_{ij} = \begin{cases} k, & \text{if } Y_{ij}(k) \geq \theta_{ij}(k - 1); \\ T_{ij}(k - 1), & \text{otherwise}. \end{cases}$$  \hspace{1cm} (22)

FIGURE 4. Results of MSMG. (a) MSMG of Fig. 2(a1); (b) MSMG of Fig. 2(b1).

FIGURE 5. Classical PCNN model.
parameter to determine the number of iterations. \( Y_{ij}(k) \) is the \( k \)-th output of PCNN.

3) MSMG-PCNN MODEL
Although the PCNN model plays an important role in image fusion, it has a critical disadvantage: For an image, one pixel corresponds one neuron in PCNN. Once PCNN is alone used, it will be very difficult to find edge information. In this situation, the fusion result may produce local block and blurring. Since the MSMG operator directly affects the edge extraction in multiple scales, it is a good choice to use the MSMG-image modulate the PCNN model. In this paper, the MSMG operator is used to adjust the linking strength \( \beta_{ij}^1 \) and \( \beta_{ij}^2 \). The expression is denoted as

\[
\beta_{ij}^1 = M_1 \\
\beta_{ij}^2 = M_2
\]

where \( M_1 \) and \( M_2 \) are the MSMG of input images \( I_1 \) and \( I_2 \), which can be computed by Formula (12). The schematic diagram of MSMG-PCNN model is shown in Fig. 6. Different from Fig. 5, the linking strengths \( \beta_{ij}^1 \) and \( \beta_{ij}^2 \) are modulated by MSMG operator.

C. CoF-MSMG-PCNN BASED FUSION METHOD
Since the MS and PAN images are acquired by different sensors, some positional deviation cannot be ignored. Image registration is required before fusion [24]. This paper focuses on the fusion algorithm. Therefore, the MS and PAN images used are all registered.

The fusion image can be divided into the following steps:

(1) The PAN image is decomposed into three sub-images \( S, L, B \), through the CoF-based decomposition method. This step can be expressed by Formulas (7)-(11), where the input image is PAN.

(2) The MS image is transformed in the HSI domain by an HSI transformation. The intensity image \( I \) will be used in the next step.

(3) A fused approximate image \( FA \) is obtained through the MSMG-PCNN model for \( B \) and \( I \). The two input images of the PCNN are \( B \) and \( I \), and the output is \( FA \).

(4) The fused intensity image \( FI \) is reconstructed by summation of \( S, L \) and \( FA \).

(5) An inverse HSI transform is applied to the \( FI \) to obtain the fused image \( F \).

It should be noted that MS and F are not only three bands. The reason for displaying RGB images is to look good in
visual effects. The diagram of the CoF-MSMG-PCNN based algorithm is shown in Fig. 7.

III. EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed algorithm, several sets of experiments are used to demonstrate. Meanwhile, several state-of-the-art algorithms, e.g., P + XS [25], HSI [6], PCA [7], guided image filtering (GIF) [26], convolutional neural network (CNN) [27], and convolutional autoencoder (CAE) [28] are used for qualitative and quantitative evaluation.

A. DATASETS

A dataset containing 27 image pairs is used in this paper. These images are captured by Landsat Enhanced Thematic
Mapper Plus (LANDSAT 7 ETM+). They can be easily acquired from U.S. Geological service [29]. Meanwhile, the data can be also acquired from [30]. The images of the dataset have a size of $400 \times 400$ pixels, which works in 6 bands (Red, Green, Blue, Near-IR, Mid-IR, and shortwave-IR). The spatial resolution of MS images is 30 m, and 15 m for PAN image. Therefore, in this paper, the MS images are subsampled into $200 \times 200$ pixels. The MS image ($200 \times 200$ pixels) and PAN image ($400 \times 400$ pixels) are merged into an MS image with $400 \times 400$ pixels. And the $400 \times 400$ pixels source MS images will be used as reference images. Fig. 8 is the display result of four randomly selected image pairs of the dataset, where (a1, b1, c1, d1) represent the size of $400 \times 400$ pixels MS images, (a2, b2, c2, d2) represent the size of $200 \times 200$ pixels MS images, and (a3, b3, c3, d3) represent the size of $400 \times 400$ pixels PAN images.

\[ \text{ERGAS}(Z, \hat{Z}) = 100 \times \frac{s_p}{s_M} \prod_{a=1}^{m_a} \left( \frac{\sqrt{\text{MSE}(Z_a, \hat{Z}_a)}}{\mu_{\hat{Z}_a}} \right) \]  

where $s_p/s_M$ is the ratio between the pixel sizes of the MS image and PAN image; $Z_a$ and $\hat{Z}_a$ represent the $a$-th bands of the fused image and the reference image; $m_a$ is the number of MS bands; $\text{MSE}(\cdot)$ is the mean square error operator, and $\mu_{\hat{Z}_a}$ is the mean of $\hat{Z}_a$. The smaller is the ERGAS, the smaller is the error in the fused image.

\[ \text{DD}(Z, \hat{Z}) = 100 \times \frac{s_p}{s_M} \prod_{a=1}^{m_a} \left( \frac{\mu_{Z_a}}{\mu_{\hat{Z}_a}} \right) \]  

where $\mu_{Z_a}$ and $\mu_{\hat{Z}_a}$ are the means of $Z_a$ and $\hat{Z}_a$, respectively. The smaller is the DD, the better is the quality of the fused image.

\[ \text{UIQI}(Z, \hat{Z}) = \prod_{a=1}^{m_a} \left( \frac{\mu_{Z_a}}{\mu_{\hat{Z}_a}} \right) \]  

where $\mu_{Z_a}$ and $\mu_{\hat{Z}_a}$ are the means of $Z_a$ and $\hat{Z}_a$, respectively. The larger is the UIQI, the better is the quality of the fused image.

**B. QUALITATIVE EVALUATION**

Since human and machine decisions are heavily dependent on visual observation, the visual effect is a very important evaluation indicator. The visual effect largely determines the quality of the fusion results. Therefore, visual evaluation is essential for qualitative evaluation.

Figs. 9-12 show four sets of experiments. All the (a) and (b) are the source MS images and PAN images, (c)-(i) are the fused results of P + XS, HSI, PCA, GIF, CNN, CAE, and the proposed CoF-MSMG-PCNN methods, respectively. Every set of experiments will be described as follow.

Fig. 9 shows the first set of experiments with different methods. It can be seen that (e) and (f) are more blurred than the MS image. (c) and (d) have similar visual performance and they look sharper than the MS image. Moreover, (g), (h), and (i) appear to be enhanced over other images, which means that further quantitative evaluation is required.

Fig. 10 shows the second set of experiments. Spectral information distortion appears in (c) and (d). (e) loses some spectral information. (f) is blurred. (g), (h), and (i) have the same situation as Fig. 9.

Fig. 11 shows the third set of experiments. Similar to Fig. 10, spectral information distortion (c) and (g) are also blurred, (e) looks darker than the MS image. (g), (h), and (i) perform better than other methods and need further quantitative verification.

Fig. 12 shows the fourth set of experiments. Unlike previous experiments, all methods show similar results and similar visual quality. From the above four sets of experiments, quantitative evaluation is necessary.

It should be noted that experiments have been performed on the entire of the dataset. Figs. 9-12 are examples of comparison experiments. The fusion results of all samples and their statistical values will be given below.

**C. QUANTITATIVE EVALUATION**

Evaluating fusion effects based on visual effects alone is subjective. Moreover, the visual effects of the CNN and CAE methods are similar to the proposed method. Therefore, it still needs enough quantitative evaluation to demonstrate the proposed method is the best.

In this section, three metrics, the relative dimensionless global error in synthesis (ERGAS), the degree of distortion (DD), and the universal image quality index (UIQI) are employed to evaluate the fused results.

(1) ERGAS: The ERGAS measures the amount of spectral distortion. Given the fused image $Z$, and the reference image $\hat{Z}$, the ERGAS is defined as

\[ \text{ERGAS}(Z, \hat{Z}) = 100 \times \frac{s_p}{s_M} \prod_{a=1}^{m_a} \left( \frac{\text{MSE}(Z_a, \hat{Z}_a)}{\mu_{\hat{Z}_a}} \right) \]  

where $s_p/s_M$ is the ratio between the pixel sizes of the MS image and PAN image; $Z_a$ and $\hat{Z}_a$ represent the $a$-th bands of the fused image and the reference image; $m_a$ is the number of MS bands; $\text{MSE}(\cdot)$ is the mean square error operator, and $\mu_{\hat{Z}_a}$ is the mean of $\hat{Z}_a$. The smaller is the ERGAS, the smaller is the error in the fused image.

(2) DD: As the name of DD, DD measures the degree of distortion between the fused image and the reference image,
TABLE 1. Quantitative evaluation of different methods.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Index</th>
<th>AWLP</th>
<th>Brovey</th>
<th>IHS</th>
<th>P+XS</th>
<th>GIF</th>
<th>PCA</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ERGAS</td>
<td>25.0416</td>
<td>24.2034</td>
<td>27.5766</td>
<td>22.6624</td>
<td>18.7522</td>
<td>19.0347</td>
<td><strong>18.6223</strong></td>
</tr>
<tr>
<td></td>
<td>DD</td>
<td>0.4317</td>
<td>0.3752</td>
<td>0.0888</td>
<td>0.0340</td>
<td><strong>0.0178</strong></td>
<td>0.0198</td>
<td>0.0185</td>
</tr>
<tr>
<td></td>
<td>UIQI</td>
<td>0.7639</td>
<td>0.8182</td>
<td>0.7459</td>
<td>0.7971</td>
<td>0.8299</td>
<td>0.8216</td>
<td><strong>0.8339</strong></td>
</tr>
<tr>
<td>2</td>
<td>ERGAS</td>
<td>28.3726</td>
<td>44.4873</td>
<td>41.0244</td>
<td><strong>23.0877</strong></td>
<td>26.3195</td>
<td>24.9346</td>
<td>25.9786</td>
</tr>
<tr>
<td></td>
<td>DD</td>
<td>0.6521</td>
<td>0.1200</td>
<td>0.1113</td>
<td>0.0566</td>
<td>0.0365</td>
<td>0.0261</td>
<td><strong>0.0234</strong></td>
</tr>
<tr>
<td></td>
<td>UIQI</td>
<td>0.8260</td>
<td>0.4291</td>
<td>0.4210</td>
<td>0.8424</td>
<td>0.8468</td>
<td>0.8397</td>
<td><strong>0.8494</strong></td>
</tr>
<tr>
<td>3</td>
<td>ERGAS</td>
<td>29.0752</td>
<td>52.8293</td>
<td>47.3452</td>
<td>29.2526</td>
<td>26.6563</td>
<td><strong>25.9177</strong></td>
<td>26.6116</td>
</tr>
<tr>
<td></td>
<td>DD</td>
<td>0.7567</td>
<td>0.1397</td>
<td>0.1237</td>
<td>0.0590</td>
<td>0.0312</td>
<td>0.0294</td>
<td><strong>0.0239</strong></td>
</tr>
<tr>
<td></td>
<td>UIQI</td>
<td>0.7686</td>
<td>0.2539</td>
<td>0.2612</td>
<td>0.8305</td>
<td>0.8383</td>
<td>0.8334</td>
<td><strong>0.8404</strong></td>
</tr>
<tr>
<td>4</td>
<td>ERGAS</td>
<td>31.5069</td>
<td>35.4873</td>
<td>29.5563</td>
<td>29.2266</td>
<td>28.0308</td>
<td>27.6358</td>
<td><strong>27.1590</strong></td>
</tr>
<tr>
<td></td>
<td>DD</td>
<td>0.4732</td>
<td>0.5160</td>
<td>0.0566</td>
<td>0.0379</td>
<td>0.0231</td>
<td>0.0302</td>
<td><strong>0.0202</strong></td>
</tr>
<tr>
<td></td>
<td>UIQI</td>
<td><strong>0.8734</strong></td>
<td>0.7879</td>
<td>0.6948</td>
<td>0.8374</td>
<td>0.8474</td>
<td>0.8403</td>
<td>0.8462</td>
</tr>
</tbody>
</table>

FIGURE 14. The curves between \( t \) and three metrics. (a) \( t \)-ERGAS; (b) \( t \)-DD; (c) \( t \)-UIQI.

Table 1 shows the results with different methods. The best performance value of every line is displayed in bold. It can be seen that the proposed method works best in most of the cases. Therefore, our proposed method is better than other methods.

TABLE 2. The mean value and standard deviation of all the fusion results.

<table>
<thead>
<tr>
<th></th>
<th>ERGAS</th>
<th>DD</th>
<th>UIQI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>26.6776</td>
<td>0.0200</td>
<td>0.7477</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>5.2061</td>
<td>0.0030</td>
<td>0.094</td>
</tr>
</tbody>
</table>

Moreover, the fusion results of all samples in the dataset is shown in Fig. 13. The mean value and standard deviation of all the fusion results are shown in Table 2.

IV. PARAMETER DISCUSSION

It should be noted that the value of parameter \( t \) in Formula (12) has a great influence on the fused result. It is because that \( t \) determines the gradient features information extracted from the input image. Fig. 14 shows the curves between \( t \) and three metrics. (a) \( t \)-ERGAS; (b) \( t \)-DD; (c) \( t \)-UIQI. It can be seen that all the three metrics will be converged with the increase of \( t \). And when \( t \) values 11, most of the datasets will acquire the best performance. Therefore, \( t \) values 11 in this paper.

\[
\text{DD}(Z, \hat{Z}) = \frac{1}{2M_m} \| \text{vec}(Z) - \text{vec}(\hat{Z}) \|_1
\]

where \( \text{vec}(\cdot) \) represents the vectorization operator, and \( \| \cdot \|_1 \) is \( l_1 \) norm.

(3) UIQI: The UIQI is used to measure image distortion, correlation and contrast between the fused image and the reference image. The UIQI is defined as

\[
\text{UIQI}(Z, \hat{Z}) = \frac{4\sigma_Z \mu_Z \cdot \mu_\hat{Z}}{(\sigma_Z^2 + \sigma_\hat{Z}^2)(\mu_Z^2 + \mu_\hat{Z}^2)}
\]

where \( \mu_Z \) and \( \mu_\hat{Z} \) are mean of \( Z \) and \( \hat{Z} \), \( \sigma_Z^2 \) and \( \sigma_\hat{Z}^2 \) are the corresponding variances, and \( \sigma_Z \hat{Z} \) is the covariance between \( Z \) and \( \hat{Z} \). The larger is the UIQI, the better is the fusion result.

The three metrics are obtained by Formulas (25)-(27). Table 1 shows the results with different methods. The best performance value of every line is displayed in bold. It can be seen that the proposed method works best in most of the cases. Therefore, our proposed method is better than other methods.
V. CONCLUSION
This paper proposes a novel remote sensing image fusion method. It is focused on the fusion of the MS image and PAN image. In this method, a very effective EPF called CoF is employed to decompose the PAN image, and a well-designed PCNN fusion strategy modulated by the MSMG operator is used to merge each band with small-scale PAN image. Finally, each fused bands are combined to obtain the fused image. Experimental results demonstrate that the proposed CoF-MSMG-PCNN method outperforms most methods.

ACKNOWLEDGMENT
The authors are grateful to the editors and the reviewers for their valuable comments and suggestions.

REFERENCES
HUIXIN ZHOU received the B.S., M.S., and Ph.D. degrees in physics electronics from Xidian University, in 1996, 2002, and 2004, respectively. He is currently a Professor with the School of Physics and Optoelectronic Engineering, Xidian University. His research interests include optoelectronic imaging, image processing, object detection, hyperspectral imaging, and machine learning.

HANLIN QIN received the B.S., M.S., and Ph.D. degrees in physics electronics from Xidian University, in 2004, 2006, and 2010, respectively. He is currently an Associate Professor with the School of Physics and Optoelectronic Engineering, Xidian University. His research interests include optoelectronic imaging, object track, hyperspectral imaging, and deep learning.