

Received October 28, 2021, accepted November 21, 2021, date of publication November 29, 2021, date of current version December 7, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3131268

Remote Sensing Image Fusion Based on Nonnegative Dictionary Learning

JUNHUA WANG¹, YUAN JIANG², YONGJUN QI³, AND YONGPING ZHAI⁴

¹Institute of Engineering, Guangzhou College of Technology and Business, Guangzhou, Guangdong 510850, China

²Ideological and Political Theory Teaching Department, South China Business College, Guangdong University of Foreign Studies, Guangzhou, Guangdong 510545, China

³Faculty of Megadate and Computing, Guangdong Baiyun University, Guangzhou, Guangdong 510450, China

⁴College of Advanced Interdisciplinary Studies, National University of Defense Technology, Changsha, Hunan 410073, China

Corresponding author: Yuan Jiang (gwng_jy@163.com)

This work was supported in part by the Major Scientific Research Projects in Guangdong Province under Grant 2018KTSCX331 and Grant 2018KQNCX378, in part by the Ministry of Education Cooperative Education under Project 201802123151 and Project 201902084029, and in part by the Feature Innovation Project of Colleges and Universities in Guangdong Province under Grant 2020ktsxcx163.

ABSTRACT For the problem of Panchromatic and multispectral remote sensing image fusion, we propose a remote sensing image fusion algorithm based on nonnegative dictionary learning. The basic idea of the algorithm is to use the panchromatic image with high spatial resolution to learn the high-low resolution dictionary pair, and to improve the fusion effect of remote sensing image by combining the nonnegativity of the image. Firstly, high resolution dictionary and low resolution dictionary are learned from high spatial resolution panchromatic image by nonnegative dictionary learning technology. Then multispectral image is sparsely represented by low resolution dictionary to obtain coefficient matrix. Finally, using coefficient matrix and high resolution dictionary, high resolution multispectral image is reconstructed. Compared with state-of-the-art methods, the proposed algorithm can get high spatial resolution and well preserve spectral information of multispectral image. Our experimental results of real QUICKBIRD and IKONOS remote sensing image fusion validate the effectiveness of the proposed algorithm.

INDEX TERMS Image fusion, pan-sharpening, dictionary learning, sparse representation, remote sensing image.


I. INTRODUCTION

Most remote sensing images obtained by optical earth observation satellites (such as IKONOS, QuickBird and Landsat) include panchromatic images with high spatial resolution and multispectral images with low spatial resolution. Panchromatic and multispectral remote sensing image fusion is also called pan sharpening. Its purpose is to obtain multispectral images with high spatial resolution and spectral resolution, and to provide detailed and accurate image data for subsequent applications such as ground object recognition and classification.

Traditional methods of remote sensing image fusion can be roughly divided into two categories: the fusion method based on component alternative and the fusion method based on multi-resolution analysis. The fusion method based on component alternative mainly includes IHS (intensity-hue-saturation) transformation method [1], PCA (principal

component analysis) principal component analysis method [2], [3] and GS (gram-Schmidt) method [4], [5]. Basic idea of this methods is to use panchromatic image to replace the components in its multispectral image and improve the spatial resolution of multispectral image. The fusion method based on multi-resolution analysis includes wavelet transform, pyramid transform and NSCT transform [6]–[8]. Its basic idea is to decompose multispectral image and panchromatic image respectively, then fuse the decomposition coefficient, and then reconstruct the fusion image through multi-resolution. This kind of algorithm can retain the spectral information in the multispectral image well, but there will be spatial distortion, typical phenomena are ringing effect, virtual scene confusion, edge and texture blur.

In recent years, with the development of sparse representation theory [9]–[12], this technology has been used in the fusion of Panchromatic and multispectral remote sensing images, and achieved advanced fusion effect. In [13], a sparse representation based remote sensing image fusion algorithm is first proposed, which uses the similar

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaojun Steven Li .

high-resolution multispectral image to learn the high-resolution dictionary and realize the high-resolution reconstruction of the low-resolution multispectral image. In [14]–[17], the low-resolution multispectral image dictionary and the high-resolution panchromatic image dictionary are obtained by using the joint dictionary. Then, the two dictionaries are used to estimate the high-resolution multispectral dictionary through different models to realize remote sensing image fusion. In [18], [19], arsis format is introduced into sparse representation image fusion algorithm to further improve the performance of the algorithm. Most of the above algorithms need to use the similar high-resolution multispectral image samples to train the dictionary, but the similar high-resolution multispectral image is often difficult to obtain, which limits the application of the algorithm in practice. In addition, with the rise of deep learning technology [20]–[22], this technology is also used in remote sensing image fusion [23]–[25], but this kind of algorithm needs a large number of samples for training.

In view of the shortcomings of existing sparse representation based on image fusion algorithms, we propose a remote sensing image fusion algorithm based on nonnegative dictionary learning. Firstly, high-low resolution dictionary pair are obtained from panchromatic images by multi task dictionary learning technology. Then the sparse representation coefficient matrix of multispectral image in low resolution dictionary is solved. Finally, the high-resolution multispectral image is reconstructed by coefficient matrix and high-resolution dictionary to realize remote sensing image fusion. Compared with the existing similar algorithms, this algorithm does not need high-resolution multispectral image as a sample, and because of the nonnegativity of image data, this algorithm can obtain the fusion image with higher spatial resolution and less spectral distortion.

The rest of this paper is organized as follows. Section II briefly describes the sparse representation and dictionary learning technology, then in Section III, the fusion algorithm of remote sensing image based on nonnegative dictionary learning is introduced in detail and Section IV gives the detailed steps of the algorithm and analyzes the computational complexity of the algorithm. In Section V, the performance of the algorithm is verified by real data experiments. Finally, conclusion is drawn in Section VI.

II. SPARSE REPRESENTATION AND DICTIONARY LEARNING

We firstly review sparse representation technology. Given a set of samples $\beta = (\beta_1, \dots, \beta_K) \in \mathbb{R}^{n \times K}$ and redundant dictionaries $A \in \mathbb{R}^{n \times m}$, the purpose of sparse representation is to find sparse coefficient matrix, so that each sample can be sparse represented. Using 1 norm represent sparsity, the problem of sparse representation is expressed as

$$\min_{A, \Gamma} \frac{1}{2} \|\beta - A\Gamma\|_F^2 + \lambda \|\Gamma\|_1 \quad (1)$$

where $\Gamma = (\Gamma_1, \Gamma_2, \dots, \Gamma_K) \in \mathbb{R}^{m \times K}$ is sparse coefficient matrix and λ is the weighted parameter. The common methods to solve sparse representation problems are the Orthogonal Matching Pursuit (OMP) and BPDN algorithm.

We then review the dictionary learning technology. Given a set of samples $\beta = (\beta_1, \dots, \beta_K) \in \mathbb{R}^{n \times K}$, the purpose of dictionary learning is to find redundant dictionaries $A \in \mathbb{R}^{n \times m}$, so that each sample can be sparse represented by sparse coefficient matrix $\Gamma = (\Gamma_1, \Gamma_2, \dots, \Gamma_K) \in \mathbb{R}^{m \times K}$. The problem of dictionary learning is expressed as

$$\min_{A, \Gamma} \frac{1}{2} \|\beta - A\Gamma\|_F^2 + \lambda \|\Gamma\|_1 \quad (2)$$

where λ is weighted parameter. The common methods to solve dictionary learning problems are mod algorithm, K-SVD algorithm [26] and online dictionary learning algorithm [27].

The MOD algorithm needs a large computational cost, so we use the idea of K-SVD algorithm to dictionary learning. K-SVD algorithm divides the process of dictionary learning into two steps: one is to update redundant dictionary $A \in \mathbb{R}^{n \times m}$ with the fixed parse coefficient matrix and the other is to update the coefficients $\Gamma = (\Gamma_1, \Gamma_2, \dots, \Gamma_K) \in \mathbb{R}^{m \times K}$ with the fixed redundant dictionary.

III. REMOTE SENSING IMAGE FUSION ALGORITHM BASED ON NONNEGATIVE DICTIONARY LEARNING

In this section, we propose a remote sensing image fusion algorithm based on nonnegative dictionary learning. In our algorithm, the high-low resolution sample set is generated by the high resolution panchromatic remote sensing image, then the high-resolution and low-resolution dictionaries are learned by using nonnegative dictionary learning technology, and the sparse representation coefficient matrix of multispectral images under the low-resolution dictionary is solved. Finally, the high-resolution multispectral images are reconstructed by using the results to realize remote sensing image fusion.

A. ABBREVIATIONS AND ACRONYMS PREPROCESSING TO GET HIGH-LOW RESOLUTION PANCHROMATIC SAMPLE SET

High resolution panchromatic remote sensing images are firstly preprocessed to get high-low resolution panchromatic sample set. The low resolution panchromatic remote sensing image Z_l is obtained by fuzzy processing and down-sampling of high resolution panchromatic remote sensing image Y_h , and then the panchromatic image Z_l is restored to the original high-resolution panchromatic image of the same size by bicubic interpolation. After the above processing, the high-low resolution panchromatic images are divided into overlapped image blocks with the same size of $\sqrt{n} \times \sqrt{n}$, and then K image blocks are randomly selected to form the high-low resolution sample set $\{p_k^h, p_k^l\} \in \mathbb{R}^n (k = 1, \dots, K)$.

The high-low resolution sample set generation process is summarized as follows:

Algorithm 1 Construct High and Low Resolution Sample Set

Input: high resolution panchromatic remote sensing image Y_h ;

(1) Down-sampling and blur the image Y_h to get the low resolution image Z_l ;

(2) The image Z_l is restore to the same size as Y_h by bicubic inter-polation, denoted as \bar{Z}_l

(3) Y_h and \bar{Z}_l are divided into overlapped image blocks with the same size, K image blocks are randomly selected;

Output: the high and low resolution sample set

$$\{p_k^h, p_k^l\} \in R^n (k = 1, \dots, K).$$

B. HIGH-LOW RESOLUTION DICTIONARY PAIR LEARNING

we use the sample set $\{p_k^h, p_k^l\}$ to learn the high resolution dictionary and low resolution dictionary through the nonnegative dictionary learning technology. Let

$$P_1 = \{p_1^h, p_2^h, \dots, p_K^h\} \in R^{n \times K}$$

$$P_2 = \{p_1^l, p_2^l, \dots, p_K^l\} \in R^{n \times K}$$

the problem of learning high-low resolution dictionary pair is written as

$$\min_{D_1, \alpha} \frac{1}{2} \|P_1 - D_1 \alpha\|_F^2 + \lambda_1 \|\alpha\|_1 \quad (3)$$

$$\min_{D_2, \alpha} \frac{1}{2} \|P_2 - D_2 \alpha\|_F^2 + \lambda_2 \|\alpha\|_1 \quad (4)$$

where $D_1, D_2 \in R^{n \times m}$ represent the high resolution dictionary and the low resolution dictionary respectively, λ_1, λ_2 are the weighted coefficient, and $\|\alpha\|_1 = \sum_{i,j} |\alpha_{i,j}|$ is the sparse constraint of the coefficient matrix $\alpha \in R^m$. By combining formula (3) and (4), we can get that the dictionary learning model is

$$\min_{D_1, D_2, \alpha} \sum_{j=1}^2 \left(\frac{1}{2} \|P_j - D_j \alpha\|_F^2 + \lambda_j \|\alpha\|_1 \right) \quad (5)$$

Image data is nonnegative. According to the theory of nonnegative matrix decomposition, P_j can be decomposed into the product of two nonnegative matrices. Therefore, by adding nonnegative constraints to the dictionary and coefficient matrix in the dictionary learning model, the computational efficiency and fusion performance are improved, and the nonnegative dictionary learning model is obtained as

$$\begin{cases} \min_{D_1, D_2, \alpha} \sum_{j=1}^2 \left(\frac{1}{2} \|P_j - D_j \alpha\|_F^2 + \lambda_j \|\alpha\|_1 \right) \\ s.t. \forall j : D_j \geq 0, \alpha \geq 0 \end{cases} \quad (6)$$

The nonnegative constraints of redundant dictionary D_j ($j = 1, 2$) and coefficient matrix α reflect the nonnegativity of hyperspectral remote sensing image data.

Based on the basic idea of nonnegative matrix decomposition multiplication iterative algorithm [28], an iterative

scheme is constructed to solve the nonnegative dictionary learning model (6). There are two iterative steps: fixed redundancy dictionary D_j ($j = 1, 2$), updating coefficient matrix α ; fixed coefficient matrix α , updating redundancy dictionary D_j ($j = 1, 2$).

When the redundant dictionary D_j ($j = 1, 2$) is fixed, the nonnegative dictionary learning model (5) can be simplified to

$$\begin{cases} \min_{\alpha} \sum_{j=1}^2 \left(\frac{1}{2} \|P_j - D_j \alpha\|_F^2 + \lambda_j \|\alpha\|_1 \right) \\ s.t. \alpha \geq 0 \end{cases} \quad (7)$$

Let

$$L_{D_1, D_2}(\alpha) = \sum_{j=1}^2 \left(\frac{1}{2} \|P_j - D_j \alpha\|_F^2 + \lambda_j \|\alpha\|_1 \right)$$

From $\alpha \geq 0$, we can get

$$\begin{aligned} \frac{\partial \|P_j - D_j \alpha\|_F^2}{\partial \alpha} &= \sum_{j=1}^2 (D_j^T D_j \alpha - D_j^T P_j) \\ &= D_1^T D_1 \alpha + D_2^T D_2 \alpha - D_1^T P_1 - D_2^T P_2 \\ &= (D_1^T D_1 + D_2^T D_2) \alpha - D_1^T P_1 - D_2^T P_2 \end{aligned}$$

From $\alpha \geq 0$, then

$$\begin{aligned} \frac{\partial \lambda_j \|\alpha\|_1}{\partial \alpha} &= \sum_{j=1}^2 (\lambda_j \mathbf{1}) \\ &= (\lambda_1 + \lambda_2) \mathbf{1} \end{aligned}$$

where $\mathbf{1}$ denotes a matrix whose element values are all 1.

So, we have

$$\begin{aligned} \frac{\partial L_{D_1, D_2}(\alpha)}{\partial \alpha} &= (D_1^T D_1 + D_2^T D_2) \alpha + (\lambda_1 + \lambda_2) \mathbf{1} \\ &\quad - (D_1^T P_1 + D_2^T P_2) \end{aligned} \quad (8)$$

From

$$\frac{\partial L_{D_1, D_2}(\alpha)}{\partial \alpha} = 0$$

we can get

$$(D_1^T D_1 + D_2^T D_2) \alpha + (\lambda_1 + \lambda_2) \mathbf{1} = (D_1^T P_1 + D_2^T P_2) \quad (9)$$

Let \otimes represent multiplication of corresponding elements in matrix, then

$$\begin{aligned} \alpha \otimes ((D_1^T D_1 + D_2^T D_2) \alpha + (\lambda_1 + \lambda_2) \mathbf{1}) \\ = \alpha \otimes (D_1^T P_1 + D_2^T P_2) \end{aligned}$$

So, we can get the iterative form of coefficient matrix α as follows

$$\alpha \leftarrow \alpha \otimes (D_1^T P_1 + D_2^T P_2) \% ((D_1^T D_1 + D_2^T D_2) \alpha + (\lambda_1 + \lambda_2) \mathbf{1}) \quad (10)$$

where $\%$ denotes Division of corresponding elements in two matrices.

When the coefficient matrix α is fixed, the nonnegative dictionary learning model (6) can be simplified as

$$\begin{cases} \min_{\mathbf{D}_1, \mathbf{D}_2} \sum_{j=1}^2 \left(\frac{1}{2} \|\mathbf{P}_j - \mathbf{D}_j \alpha\|_F^2 \right) \\ s.t. \forall j: \mathbf{D}_j \geq 0 \end{cases} \quad (11)$$

Let

$$L_\alpha(\mathbf{D}_1, \mathbf{D}_2) = \sum_{j=1}^2 \left(\frac{1}{2} \|\mathbf{P}_j - \mathbf{D}_j \alpha\|_F^2 \right)$$

we can get

$$\frac{\partial L_\alpha(\mathbf{D}_1, \mathbf{D}_2)}{\partial \mathbf{D}_j} = \mathbf{D}_j \alpha \alpha^T - \mathbf{P}_j \alpha^T \quad (12)$$

Let

$$\frac{\partial L_\alpha(\mathbf{D}_1, \mathbf{D}_2)}{\partial \mathbf{D}_j} = 0$$

then

$$\mathbf{D}_j \otimes (\mathbf{D}_j \alpha \alpha^T) = \mathbf{D}_j \otimes (\mathbf{P}_j \alpha^T) \quad (13)$$

From the above formula, the iterative form of redundant dictionary can be written as

$$\mathbf{D}_j \leftarrow \mathbf{D}_j \otimes (\mathbf{P}_j \alpha^T) \% (\mathbf{D}_j \alpha \alpha^T) \quad (14)$$

Specifically, the iterative form of high resolution dictionary is denoted as

$$\mathbf{D}_1 \leftarrow \mathbf{D}_1 \otimes (\mathbf{P}_1 \alpha^T) \% (\mathbf{D}_1 \alpha \alpha^T)$$

The iterative form of low resolution dictionary is written as

$$\mathbf{D}_2 \leftarrow \mathbf{D}_2 \otimes (\mathbf{P}_2 \alpha^T) \% (\mathbf{D}_2 \alpha \alpha^T)$$

Repeat the above two iterative steps (10) and (14) until convergence, and the high-resolution dictionary \mathbf{D}_1 , low-resolution dictionary \mathbf{D}_2 and coefficient matrix α can be obtained. It can be seen from the iterative schemes (10) and (14) that when the initial values of dictionary \mathbf{D}_j ($j = 1, 2$) and coefficient matrix α are nonnegative matrices, the updated dictionary and coefficient matrix are also nonnegative matrices.

C. HIGH RESOLUTION MULTISPECTRAL REMOTE SENSING IMAGE RECONSTRUCTION

High resolution multispectral remote sensing image can be reconstructed by the high resolution dictionary and low resolution dictionary. The input low resolution multispectral image is used to estimate its sparse representation coefficient matrix under the low resolution dictionary, and then the high resolution multispectral image is reconstructed by combining the sparse representation coefficient matrix and the high resolution dictionary. Denote the low resolution multispectral images as \mathbf{x}_s ($s = 1, \dots, S$) and S represent the number of bands included. Image \mathbf{X}_s is restored to low resolution interpolating image $\tilde{\mathbf{X}}_s$ by bicubic interpolation, which is the same size as the original high resolution panchromatic image and

then $\tilde{\mathbf{X}}_s$ is divided into P image blocks which overlap each other and have the same size as $\sqrt{n} \times \sqrt{n}$. Denote the vector of the p -th image block as \mathbf{x}_p^s and the data matrix of all the s -band image blocks as $\mathbf{x}_s = [\mathbf{x}_1^s, \dots, \mathbf{x}_p^s] \in \mathbb{R}^{n \times P}$, then the nonnegative sparse representation model of the coefficient matrix is

$$\begin{cases} \min_{\omega_s} \frac{1}{2} \|\mathbf{x}_s - \mathbf{D}_2 \omega_s\|_F^2 + \lambda_2 \|\omega_s\|_1 \\ s.t. \omega_s \geq 0 \end{cases} \quad (15)$$

where $\omega_s \in \mathbb{R}^{m \times P}$ is the sparse representation coefficient matrix.

Let

$$L_{\omega_s}(\omega_s) = \frac{1}{2} \|\mathbf{x}_s - \mathbf{D}_2 \omega_s\|_F^2 + \lambda_2 \|\omega_s\|_1,$$

we can get

$$\frac{\partial L_{\omega_s}(\omega_s)}{\partial \omega_s} = \mathbf{D}_2^T \mathbf{D}_2 \omega_s - \mathbf{D}_2^T \mathbf{x}_s + \lambda_2 \mathbf{1} \quad (16)$$

From

$$\frac{\partial L_{\omega_s}(\omega_s)}{\partial \omega_s} = 0$$

equation (16) can be written as

$$\mathbf{D}_2^T \mathbf{x}_s = \mathbf{D}_2^T \mathbf{D}_2 \omega_s + \lambda_2 \mathbf{1} \quad (17)$$

Similar to the iterative scheme (10), we can get the following iterative scheme

$$\omega_s \leftarrow \omega_s \otimes (\mathbf{D}_2^T \mathbf{x}_s) \% (\mathbf{D}_2^T \mathbf{D}_2 \omega_s + \lambda_2 \mathbf{1}) \quad (18)$$

The coefficient matrix of each band image can be obtained by using the iterative scheme (18). According to the iterative scheme (18), when the initial value of the coefficient matrix ω_s is a non-negative matrix, ω_s updated by the iterative scheme is also a nonnegative matrix.

Next, we use the high resolution dictionary \mathbf{D}_1 and coefficient matrix ω_s to reconstruct the high resolution multispectral image. Using high resolution dictionary \mathbf{D}_1 and coefficient matrix ω_s , the reconstructed image data matrix of each band is

$$\mathbf{x}_s^h = \mathbf{D}_s \omega_s \quad s = 1, \dots, S \quad (19)$$

Each column vector in the data matrix \mathbf{x}_s^h corresponds to an image block in the s -band. The high-resolution multispectral remote sensing image $\tilde{\mathbf{X}}_s$ ($s = 1, \dots, S$) can be obtained by splicing the image blocks according to their positions and averaging the overlapped parts.

IV. DETAILED STEPS AND ANALYSIS OF OUR ALGORITHM

A. DETAILED STEPS OF OUR ALGORITHM

In this subsection, the detailed steps of the proposed algorithm are described, as shown in Algorithm 2.

Algorithm 2 Remote Sensing Image Fusion Algorithm Based on Nonnegative Dictionary Learning

Input: multispectral image X_s ($s = 1, \dots, S$) and panchromatic image Y_h , weighting parameters λ_1, λ_2 , number of dictionary atoms m ;

Initialization: the initial high resolution dictionary D_1 and low resolution dictionary D_2 , the elements in the initial coefficient matrix α and ω_s are random values between 0 and 1;

(1) learn high resolution dictionary D_1 and low resolution dictionary D_2

(1-1) pre-processing of panchromatic image Y_h and random sampling to obtain high-low resolution sample set $\{p_k^h, p_k^l\}$;

(1-2) using the sample set $\{p_k^h, p_k^l\}$, the high resolution dictionary D_1 and low resolution dictionary D_2 are learned;

(1-2-1) update coefficient matrix α by using iterative scheme (9);

(1-2-2) update high resolution dictionary D_1 and low resolution dictionary D_2 by using iterative scheme (13)

(1-2-3) repeat steps (1-2-1) and (1-2-2) until convergence to obtain high resolution dictionaries D_1 and low resolution dictionaries D_2 ;

(2) Estimate sparse representation coefficient matrix ω_s of multispectral image

(2-1) update the coefficient matrix ω_s by using the iterative scheme (17);

(2-2) Repeat step (2-1) until it converges to get the coefficient matrix ω_s ;

(3) use dictionary D_1 and coefficient matrix ω_s to recover the image of each band through equation (18);

Output: multispectral remote sensing image after fusion

$$\hat{X} = [\hat{X}_1, \dots, \hat{X}_S]$$

Note for Algorithm 2:

1) SELECTION OF DICTIONARY DIMENSION M

The high-low resolution dictionary pair have the same dimension, and the dictionary dimension is generally taken as $2n \leq m \leq 4n$. If the dictionary dimension is too large, the calculation amount will increase.

2) SELECTION OF THE WEIGHT PARAMETERS λ_1 AND λ_2

Parameters λ_1 and λ_2 are values generally related to dictionary dimension m , and the weight parameter of our algorithm is set as

$$\lambda_1 = \lambda_2 = \sqrt{2 \log(m)}$$

B. ANALYSIS OF COMPUTATIONAL COMPLEXITY

The computation cost of our algorithm is mainly from two steps: learning low-high resolution dictionary pair and solving low-resolution multispectral image coefficient matrix ω_s . The cost of computation for high-low resolution dictionary

pair learning comes from the coefficient matrix α iterative format (10) and the high and low resolution dictionary learning iterative format (14). Assuming that the dictionary learning iterations N times, the cost of computation required is respectively

$$T_\alpha = O(NKnm) + O(NKm^2)$$

$$T_{D_1, D_2} = O(NKm^2)$$

Because $n < m$, computation cost for learning high-low resolution dictionary is

$$T_{DL} = O(NKm^2)$$

The computation cost of coefficient matrix ω_s is derived from the iterative scheme (18). If the number of iterations is N_1 , the computation cost of the estimated coefficient matrix ω_s is expressed as

$$T_\omega = O(N_1 P m^2)$$

Because multispectral images have S bands in total, the computational cost of this algorithm is

$$T = T_{DL} + S T_\omega = O(NKm^2) + O(SN_1 P m^2) \quad (20)$$

It can be seen that the time complexity of this algorithm is related to the number of bands S , the number of iterations N and N_1 , the number of samples K , the number of image patches P and the dictionary dimension m , and is most affected by the dictionary dimension m .

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we will verify the performance of our algorithm through real data experiments, and compare it with GS algorithm [4] and SR-D algorithm [19]. Subsection A compares the performance of real QuickBird Remote sensing image after fusion; Subsection B compares the performance of real IKONOS remote sensing image after fusion; Subsection C shows the influence of dictionary dimension on algorithm performance; Subsection D shows the influence of training dictionary sample number on algorithm performance.

In the following experiments, panchromatic and multispectral images provided by the QuickBird and IKONOS satellites are used. QuickBird Remote sensing images include high-resolution panchromatic images with a resolution of 0.7 m and low-resolution multispectral images with a resolution of 2.8 m, among which multispectral image contains four bands. IKONOS remote sensing image includes high-resolution panchromatic image with a resolution of 1 meter and low-resolution multispectral image with a resolution of 4 meters, among which multispectral image contains 4 bands. The size of image block is 8×8 , the overlapping pixels of adjacent cubes are 7, the number of dictionary atoms is $m = 256$, and the number of samples is $K = 4000$.

Due to the lack of ideal fusion image as a reference, we use spectral angle mapper (SAM) index and average

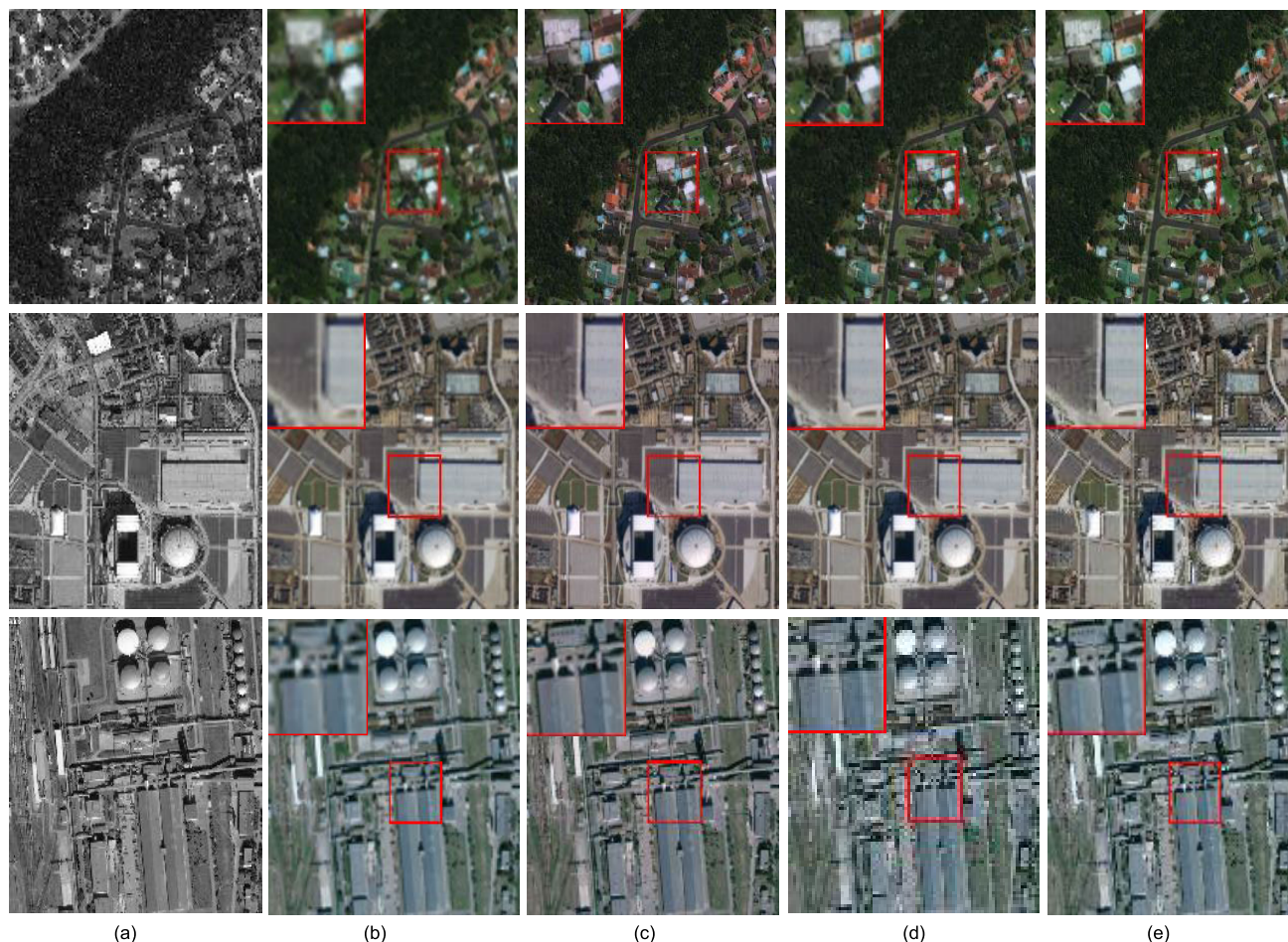


FIGURE 1. Comparison of fusion results of real QuickBird Remote Sensing Image. (a) Original panchromatic image; (b) Original multispectral image; (c) GS algorithm; (d) SR-D algorithm; (e) Our algorithm.

gradient (AG) index to measure the degree of spectral and spatial detail information retention respectively in this paper. SAM uses the original multispectral image as the reference image to measure the degree of spectral information retention of the original multispectral image. The smaller the value of SAM, the better the image spectral information retention. AG uses high-resolution panchromatic image as reference image to measure the degree of spatial detail preservation. The higher the value of AG, the better the image spatial detail retention.

A. COMPARISON OF FUSION PERFORMANCE OF REAL QUICKBIRD REMOTE SENSING IMAGE

In this subsection, we compare the performance of three algorithms by fusing the panchromatic image and multispectral image of QuickBird satellite. Figure 1 compares the visual effects of three QuickBird Remote sensing images after fusion. As can be seen from Figure 1, compared with GS algorithm and SR-D algorithm, the image color fused by this algorithm is closer to the original multispectral image, which can maintain better spectral information, and the image details are clearer. Table 1 compares the SAM index and

AG index of the three algorithms of QuickBird Remote Sensing Image (the mean value of the three fusion images). As can be seen from Table 1, our algorithm has the minimum SAM value and the maximum AG value. Compared with the traditional GS algorithm, the value of SAM algorithm is about 9 less. Compared with the similar SR-D algorithm, the value of SAM algorithm is about 1.2 less. This experiment shows that our algorithm has better visual effect, and that our maintains better spectral information and higher spatial detail information.

B. COMPARISON OF FUSION PERFORMANCE OF REAL IKONOS REMOTE SENSING IMAGE

Through the fusion of IKONOS satellite panchromatic image and multispectral image, this experiment compares the performance of three algorithms. Figure 2 compares the visual effects of three IKONOS remote sensing images after fusion. As can be seen from Figure 2, compared with GS algorithm and SR-D algorithm, the color of the image fused by this algorithm is closer to the original multispectral image Figure 2 (b), and the details of the image are clearer. Table 2 compares the SAM indicators and Ag indicators of the three algorithms.

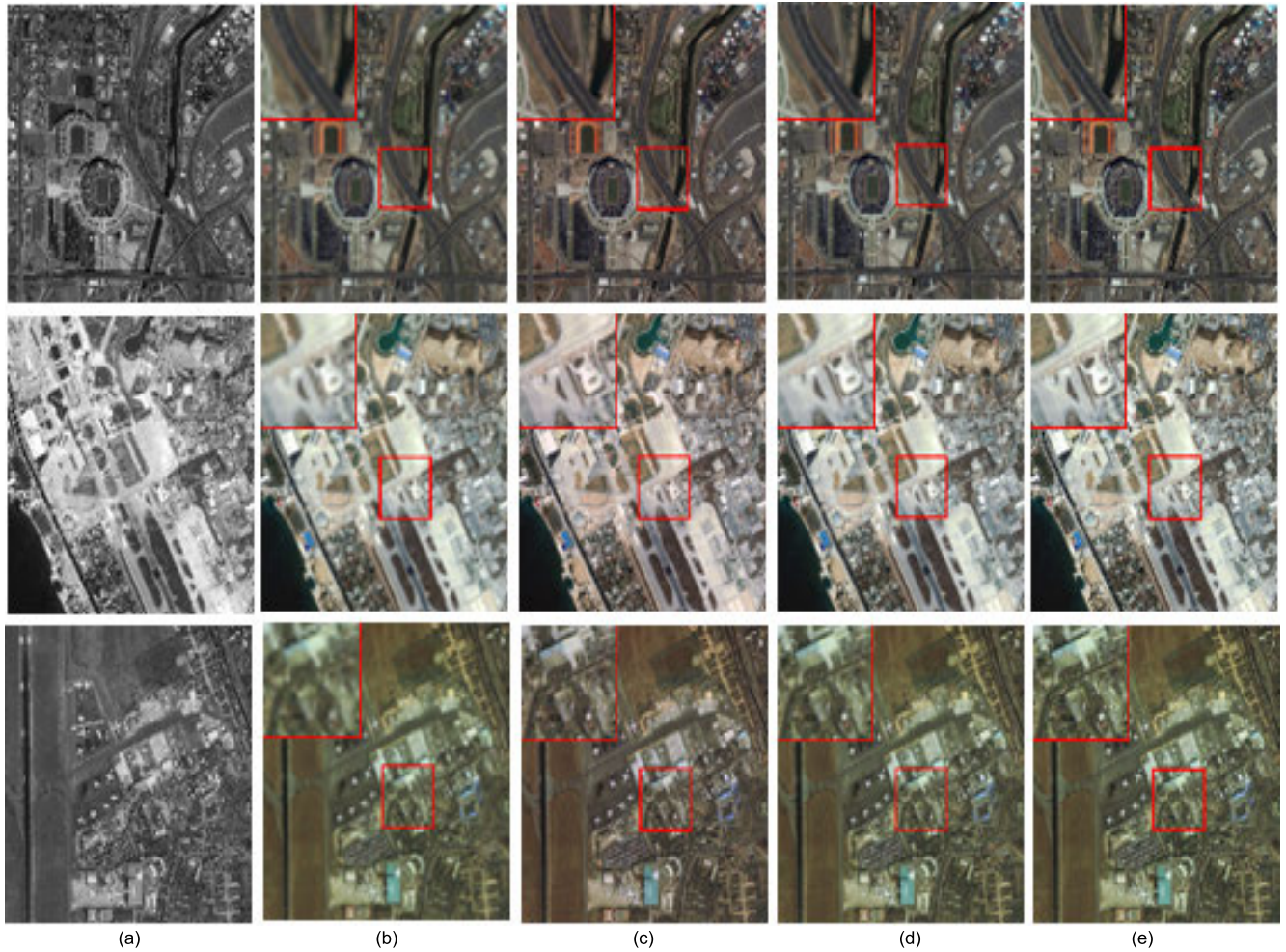


FIGURE 2. Comparison of real IKONOS remote sensing image fusion results. (a) Original panchromatic image; (b) Original multispectral image; (c) GS algorithm; (d) SR-D algorithm; (e) Our algorithm.

TABLE 1. Compare SAM and AG indexes for quickbird images.

index	GS algorithm	SR-D algorithm	Our algorithm
SAM	12.1981	4.5734	3.2604
AG	0.0278	0.0210	0.0329

As can be seen from Table 2, the algorithm in this paper has the minimum SAM value and the maximum AG value. Compared with the traditional GS algorithm, the value of SAM algorithm is about 9 less. Compared with the similar SR-D algorithm, the value of SAM algorithm is about 1 less. This experiment also shows that our algorithm has better visual effect, and that our maintains better spectral information and higher spatial detail information.

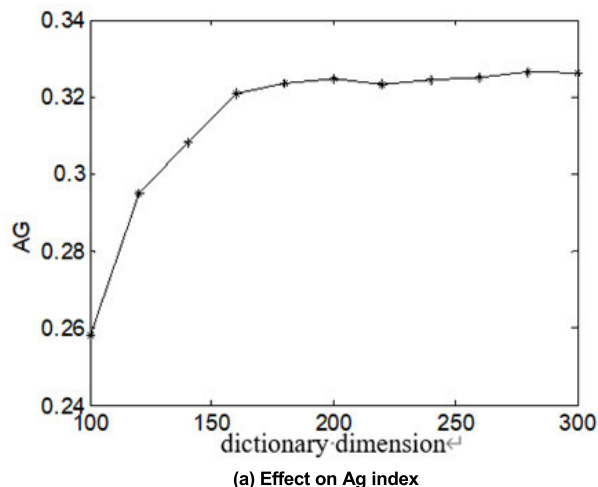
C. INFLUENCE OF DICTIONARY DIMENSION ON ALGORITHM PERFORMANCE

This experiment shows the effect of the number of samples needed to train the high-low resolution dictionary pair on the performance of our algorithm. Using three remote

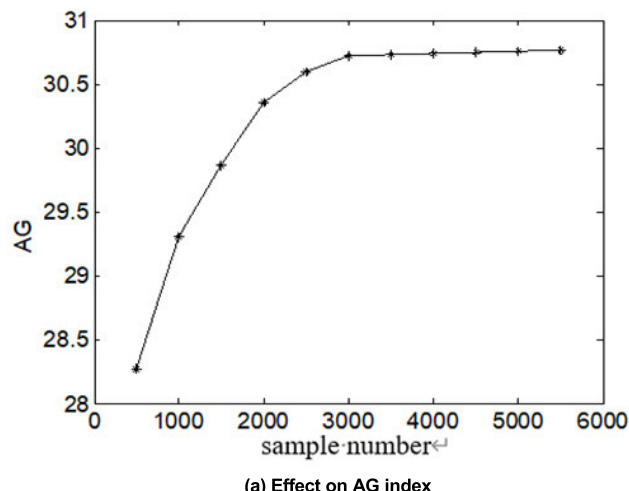
sensing images of QuickBird satellite in subsection A, the dictionary dimension (or the number of dictionary atoms) is from 100 to 300, and other simulation conditions are unchanged. Figure 3 (a) and Figure 3 (b) respectively show the SAM and AG indexes of the algorithm in this paper when the dictionary dimensions are different. It can be seen from Figure 3 that when the dictionary dimension is less than 160, the Ag value increases gradually with the increase of dictionary dimension, while the SAM value decreases gradually. When the dictionary dimension is greater than or equal to 160, the algorithm performance is stable. Therefore, in order to keep the good performance of this algorithm, the selection of dictionary dimension should not be less than 160. That is to say, the m value of dictionary $A \in \mathbb{R}^{n \times m}$ needs to be greater than 160.

D. THE EFFECT OF SAMPLE NUMBER ON ALGORITHM PERFORMANCE

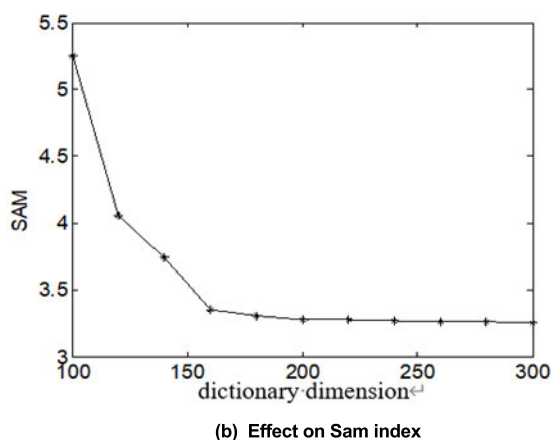
This experiment shows the effect of the number of samples needed to train the high-low resolution dictionary on the



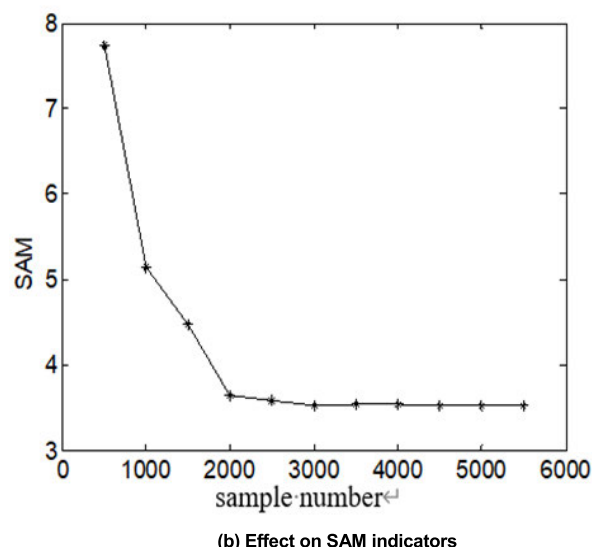
(a) Effect on Ag index



(a) Effect on AG index



(b) Effect on Sam index



(b) Effect on SAM indicators

FIGURE 3. Effect of dictionary dimension on algorithm performance.

TABLE 2. Compare SAM and AG indexes for IKONOS remote sensing image.

index	GS algorithm	SR-D algorithm	Our algorithm
SAM	12.2386	4.4707	3.3275
AG	0.0292	0.0256	0.0314

performance of the algorithm. Using the remote sensing image of Quick Bird satellite in subsection A, the number of samples changes from 500 to 5500, and other simulation conditions remain unchanged. Figure 4 (a) and Figure 4 (b) respectively show the S As can be seen from Figure 4, when the number of samples is less than 3000, the AG value increases gradually with the increase of the number of samples, while the SAM value decreases gradually. When the number of samples is greater than or equal to 3000, the performance of the algorithm is basically stable. It can be seen that when training high-low resolution dictionary pair, the number of randomly selected samples should not be less than 3000 pairs. That is to say, the K value of the sample set needs to be greater than 160. AM and AG indexes of the algorithm in this paper.

FIGURE 4. Effect of sample number on algorithm performance.

VI. CONCLUSION

To solve the fusion problem of high resolution panchromatic remote sensing image and low resolution multispectral remote sensing image, a remote sensing image fusion algorithm based on nonnegative dictionary learning is proposed. The basic idea of the algorithm is to use the high-resolution panchromatic image and the degraded low-resolution panchromatic image to learn the high-resolution and low-resolution dictionary, and to improve the fusion performance by combining the nonnegativity of the image. First, the high-resolution and low-resolution dictionaries are obtained by using the nonnegative dictionary learning technology of panchromatic images. Then the sparse representation coefficient matrix of multispectral image in low resolution dictionary is solved. Finally, the high-resolution multispectral image is reconstructed by the coefficient matrix and the high-resolution dictionary to realize the fusion of remote sensing image. Through the experiments of high-resolution panchromatic remote sensing image fusion of the QuickBird and IKONOS satellites, it is shown

that the fusion result of this algorithm can obtain higher spatial resolution and better maintain the spectral information of multispectral image.

REFERENCES

- [1] T. M. Tu, S. C. Su, H. C. Shyu, and P. S. Huang, "A new look at Ihs-like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, 2001.
- [2] V. P. Shah, N. H. Younan, and R. L. King, "An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1323–1335, May 2008.
- [3] D. Li, W. E. Wong, W. Wang, Y. Yao, and M. Chau, "Detection and mitigation of label-flipping attacks in federated learning systems with KPCA and K-means," in *Proc. 8th Int. Conf. Dependable Syst. Their Appl. (DSA)*, Yinchuan, China, 2021, pp. 551–559.
- [4] C. Laben and B. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," U.S. Patent 6011 875, Jan. 4, 2000.
- [5] Q. Xu, B. Li, and L. Ding, "High-fidelity component substitution pan-sharpening by the fitting of substitution data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 7380–7392, Nov. 2014.
- [6] X. Kang, S. Li, and J. Benediktsson, "Pansharpening with matting model," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5088–5099, Aug. 2014.
- [7] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.
- [8] J. L. Starck, J. Fadili, and F. Murtagh, "The undecimated wavelet decomposition and its reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 297–309, Feb. 2007.
- [9] M. Elad, "Sparse and redundant representation modeling—What next?" *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 922–928, Dec. 2012.
- [10] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer, 2010. [Online]. Available: <https://dl.acm.org/doi/book/10.5555/1895005>
- [11] V. M. Patel and R. Chellappa, *Sparse Representations and Compressive Sensing for Imaging and Vision*. New York, NY, USA: Springer, 2013. [Online]. Available: <https://link.springer.com/book/10.1007%2F978-1-4614-6381-8>
- [12] L. Zhang and S. Liu, "Quality evaluation of multi-frame sports image fusion based on gradient domain," *Int. J. Performability Eng.*, vol. 16, no. 3, pp. 411–418, Mar. 2020. [Online]. Available: <http://www.ijpe-online.com/EN/10.23940/ijpe.20.03.p10.411418>
- [13] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, Feb. 2011.
- [14] S. Li, H. Yin, and L. Fang, "Remote sensing image fusion via sparse representations over learned dictionaries," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4779–4789, Sep. 2013.
- [15] C. Jiang, H. Zhang, H. Shen, and L. Zhang, "A practical compressed sensing-based pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 629–633, Jul. 2012.
- [16] M. Ghahremani, Y. Liu, P. Yuen, and A. Behera, "Remote sensing image fusion via compressive sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 34–48, Jun. 2019.
- [17] H. Yin, "PAN-guided cross-resolution projection for local adaptive sparse representation-based pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4938–4950, Jul. 2019.
- [18] M. Cheng, C. Wang, and J. Li, "Sparse representation based pansharpening using trained dictionary," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 293–297, Jan. 2014.
- [19] M. R. Vicinanza, R. Restaino, G. Vivone, M. Dalla Mura, and J. Chanussot, "A pansharpening method based on the sparse representation of injected details," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 180–184, Jan. 2015.
- [20] F. Es-Sabery, A. Hair, and J. Qadir, "Sentence-level classification using parallel fuzzy deep learning classifier," *IEEE Access*, vol. 5, pp. 502–513, 2021.
- [21] M. Stuart and M. Manic, "Deep learning shared bandpass filters for resource-constrained human activity recognition," *IEEE Access*, vol. 6, pp. 307–319, 2021.
- [22] W. Xue, Q. Li, and Q. Xue, "Text detection and recognition for images of medical laboratory reports with a deep learning approach," *IEEE Access*, vol. 8, pp. 407–416, 2020.
- [23] X. Li, F. Xu, X. Lyu, Y. Tong, Z. Chen, S. Li, and D. Liu, "A remote-sensing image pan-sharpening method based on multi-scale channel attention residual network," *IEEE Access*, vol. 2020, 8, pp. 49812–49822.
- [24] T. Benzenati, Y. Kessentini, A. Kallel, and H. Hallabia, "Generalized Laplacian pyramid pan-sharpening gain injection prediction based on CNN," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 4, pp. 651–655, Apr. 2020.
- [25] Q. Wu and X. Zhai, "Remote sensing object detection via an improved Yolo network," *Int. J. Performability Eng.*, vol. 16, no. 11, pp. 1803–1813, Nov. 2020.
- [26] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [27] C. Lu, J. Shi, and J. Jia, "Online robust dictionary learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 415–422.
- [28] D. Lee and H. Seung, "Algorithms for non-negative matrix factorization," *Adv. Neural Inf. Process. Syst.*, vol. 19, no. 13, pp. 556–562, 2001.



JUNHUA WANG was born in Jingdezhen, Jiangxi, China, in 1980. He received the B.S. and M.S. degrees from Northwestern Polytechnic University, China, and the Ph.D. degree from the National University of Defense Technology. He is currently working with the Guangzhou College of Technology and Business. His research interests include computer vision, passive location, and machine learning.



YUAN JIANG was born in Huangshi, Hubei, China, in 1985. She received the B.S. degree from Changsha University and the M.S. degree from Yunnan Normal University, China. She is currently working with the South China Business College, Guangdong University of Foreign Studies. Her research interests include video analysis, teaching reform, and teaching quality evaluation.



YONGJUN QI was born in Changde, Hunan, China, in 1978. He received the B.S. and M.S. degrees from Hunan Normal University, China. He is currently working with the Faculty of Megadate and Computing, Guangdong Baiyun University. His research interests include data mining and sparse representation.



YONGPING ZHAI was born in Tianshui, Gansu, China, in 1982. He received the B.S., M.S., and Ph.D. degrees from the National University of Defense Technology. He is currently working with the College of Advanced Interdisciplinary Studies, National University of Defense Technology. His research interests include computer vision, deep learning, and medical robotics.