

Image Fusion in Remote Sensing Based on Spectral Unmixing and Improved Non-negative Matrix Factorization Algorithm

Chunlin He¹, Juan Wang^{1,*}, Siyu Lai² and A. Ennadi³

¹College of Computer Science, China West Normal University, Nanchong 637002, China

²Department of Medical Imaging, North Sichuan Medical College, Nanchong 637000, China

³School of Computer Engineering, KIIT University, Bhubaneswar, Odisha, 751024, India

Received 3 March 2018; Accepted 24 May 2018

Abstract

The spectral resolution of images improves with the continuous development of imaging technology in remote sensing. Traditional image fusion technologies result in problems in hyperspectral image fusion, such as spectral distortion and unclear physical meaning. A fast model based on improved non-negative matrix factorization (INMF) for hyperspectral and panchromatic image fusion was proposed to optimize the results of image fusion. First, the hyperspectral image was decomposed into endmember and abundance matrices using the INMF algorithm. The panchromatic image was then used to sharpen the abundance matrix. The spectral and sparse constraints were imposed on the objective function of the model. Finally, the resultant fused image was obtained by reconstructing the resolved endmember and abundance matrices. Results show that the proposed method is superior to other methods in terms of subjective assessment and objective analysis. As far as the indices of information entropy, correlation coefficient, Q -average, and spectral divergence are concerned, the proposed method surpasses those of second-best methods (Ehler, classical NMF, Ehler and Ehler) by 2.06%, 0.36%, 0.91%, and 56.31%, respectively, in the synthetic data experiment, and exceeds those of the second-best methods (Ehler, high-pass filtering, high-pass filtering and Ehler) by 0.13%, 10.05%, 3.89%, and 7.26%, respectively, in the real data experiment. Moreover, runtime is proportional to data size, and the proposed method takes the least time when image size is between 1 MB and 480 MB. This study provides theoretical reference for fast hyperspectral and panchromatic image fusion.

Keywords: Hyperspectral image fusion, Spectral unmixing, Non-negative matrix factorization, Linear spectral mixture model

1. Introduction

Image fusion integrates high spectral resolution images (HSRI) with low geometric resolution and high geometric resolution images (HGRI) with low spectral resolution [1]. Given the physical limitations of the spectrometer, the spatial resolution of HSRI cannot satisfy the requirements for practical applications. However, HSRI and HGRI can be merged in the same scene into high spectral and geometric resolution images (HSGRI).

Given the diversity of the earth's surface, some pixels in HSRI contain a variety of materials, which are called mixed pixels. From the point of spectral preservation, the integrated HSGRI should have the same endmembers as the original HSRI. However, a number of methods in image analysis for remote sensing have failed to prove that mixture theory is reasonable. Many scholars observed that spectral unmixing is a good tool for solving this problem [2][3][4]. Hence, extracting endmember and abundance coefficients from hyperspectral image of remote sensing with widely mixed pixels and producing effectively fused image are important content of the quantitative analysis of image of remote sensing.

Extensive studies were conducted in the field of image preprocessing, model building, and algorithm optimization

of hyperspectral image fusion [5][6][7][8]. However, this fusion model, which has clear physical meaning and time-efficient, cannot be easily optimized to enhance the spectral and spatial resolutions of the fused image. Therefore, the fusion method, algorithm optimization, and quantitative control of the spectral and spatial resolutions are crucial in the design of fusion model.

The present study establishes a fusion model for hyperspectral image with spectral and sparse constraints by employing a fast approach based on spectral unmixing. The results of spectral unmixing, resolution enhancement, spectral preservation, and convergence of the objective function are analyzed and compared to prove its effectivity and practicability.

2. State of the art

A series of investigations were conducted on image fusion of remote sensing. Fusion methods with simple arithmetic structure obtain HSGRI by replacing a certain component of HSRI with the sharpened HGRI. However, given that HGRI and HSRI bands are quite different in spectral signature, these methods often lead to color distortion [9][10][11][12]. Other fusion methods transform the fusion process into the optimization of a specific function under a certain statistical model [13][14][15][16]. These methods are theoretically clear, but their framework is complex for estimation and the

*E-mail address: wjuan0712@126.com

ISSN: 1791-2377 © 2018 Eastern Macedonia and Thrace Institute of Technology. All rights reserved.

doi:10.25103/jestr.113.11

occurrence of incomplete deconvolution may result in an inefficient model. Multiscale and multiresolution methods extract high-frequency parts from HGRI and inject them into HSRI by decomposing the source image in a specific domain [17][18][19][20][21]. However, this type of method is highly dependent on the spectral properties of sampling methods and fails to intensify the spatial resolution of all bands.

Fusion algorithms based on the spectral unmixing model were also widely used. Halimi et al. proposed a bilinear model that introduced an additive interaction term. Given that certain specific a priori relative to the endmember is required in advance, the bilinear model hinders the application in dealing with multi-material mixed data [22]. Xu et al. investigated the orthogonal principal component analysis (PCA) method based on second-order decorrelation, singular value decomposition (SVD) method based on projection pursuit, and independent component analysis (ICA) approach, which relies on the statistical feature of high-order signal. Their findings show that methods based on PCA, SVD, or ICA share the same problem, wherein the extracted endmembers cannot be guaranteed to be non-negative and the physical meaning of the estimated abundance is weakly correlated with the original data [23]. Kubokawa et al. considered each vertex of a simplex as an endmember under the theory of convex geometry. However, the recognition of convex hull under this method is computationally time-consuming and requires pure pixel assumption [24]. Hoyer presented a sparsely constrained non-negative matrix factorization (NMF) method, which quantitatively achieved any desired degree of sparseness by applying L_1 - and L_2 -norms; however, this method failed to explain how to set parameters adaptively [25]. Jia et al. introduced piecewise smooth and sparse constraints into NMF, which makes NMF robust to negative factors; however, the optional degree of sparseness may cause uncertainty in unmixing [26]. Yokoya et al. [27] defined a method based on coupled non-negative matrix factorization (CNMF), which is characterized as a dual-loop structure that comprises updating and unmixing operation to generate cost functions with respect to spectral and spatial degradation convergence. Nevertheless, the practical application of CNMF should be proven because of the difficulty of involving and implementing numerous influencing factors [28].

These problems mainly lie in color distortion, non-negative result, assumption of pure pixels, and insufficient convergence. This study proposes an improved non-negative matrix factorization (INMF) algorithm, which has sparse and spectral constraints and is applicable to hyperspectral and panchromatic image fusion. The fusion model for hyperspectral and panchromatic images is established through the proposed algorithm. The cost function of the traditional NMF method is optimized and runtime is reduced on the basis of the analysis of the characteristics of the linear spectral mixture model (LSMM). The sparse constraint forced on the abundance matrix is discussed, as well as the term of the spectral angle that functions as a measure for estimating the spectral consistency of HSGRI and HSRI. Furthermore, the effect of the proposed model on spectral preservation and spatial enhancement is verified to provide helpful reference for the establishment of the hyperspectral fusion model.

The remainder of this study is organized as follows. Section 3 introduces the LSMM, illustrates the workflow of the INMF algorithm, and describes the implementation of the proposed model. Section 4 visually and objectively compares the proposed method with typical methods and

reports the computational efficiency of the proposed algorithm. Section 5 summarizes the conclusions.

3. Methodology

3.1 Linear spectral mixture model

Linear spectral mixing is the most commonly used data model in image analysis for hyperspectral remote sensing. In this model, the observed value of any pixel is supposed to be the linear combination of the spectral signature and the abundance fraction of the relevant endmember. Given its rational physical significance, the LSMM is extensively used in hyperspectral image processing, which is formulated as follows:

$$R = MS + E, \quad (1)$$

where $R \in \mathbb{R}^{L \times N}$ is the image cube and L and N denote for the number of spectral bands and the number of pixels of the image, respectively. $M \in \mathbb{R}^{L \times P}$ is the matrix of the spectral signature, with each column vector $\{M_j\}_{j=1}^P \in \mathbb{R}^L$ representing an endmember spectrum, and P is the number of endmembers in the image. $S \in \mathbb{R}^{P \times N}$ is the abundance matrix. Each column vector $\{S_k\}_{k=1}^N \in \mathbb{R}^P$ denotes the abundance fraction of endmembers for the k^{th} pixel, and $E \in \mathbb{R}^{L \times N}$ can be interpreted as noise. According to the physical significance of the LSMM, the spectral signatures and abundance fractions must be non-negative, the abundance of any endmember at any pixel must be constrained from 0 to 1, and all of the fractions for each pixel should sum to unity. These conditions can be formalized as follows:

$$M_{lp} \geq 0, 1 \leq l \leq L \text{ and } S_{np} \in [0, 1], \sum_{p=1}^P S_{np} = 1, 1 \leq n \leq N \quad (2)$$

3.2 Fast INMF algorithm

NMF was first proposed by Paatero and Tapper [29]. NMF has drawn considerable attention and has been employed in widespread applications because of its simplicity and flexible interpretation. For a classical NMF problem, NMF searches two non-negative factors $W \in \mathbb{R}^{m \times k}$ and $H \in \mathbb{R}^{k \times n}$ that provide a lower rank approximation of V given a non-negative matrix $V \in \mathbb{R}^{m \times n}$ and an appropriate rank $k \ll \min(m, n)$, which satisfies the following condition:

$$V \approx WH \text{ s.t. } W, H \geq 0. \quad (3)$$

Given that Formulas (3) and (1) have a similar mathematical form and are constrained by the same terms by applying the NMF solution to the issue of spectral unmixing is natural. In finding the desired factors W and H , the traditional NMF tries to transform Equation (3) into a reformulated form, as shown in Equation (4), which is called the Euclid distance function; $\|\cdot\|$ represents the Frobenius norm:

$$\min_{W, H} f(W, H) = \frac{1}{2} \sum_{j=1}^k \|V_j - WH_j\|_2^2 = \frac{1}{2} \|V - WH\|_F^2, \text{ s.t. } W, H \geq 0 \quad (4)$$

Chen [30] proposed a NMF-based algorithm. Each entry of W and H is constrained to be no less than a minimum value of positive ε rather than 0 to make the quadratic coefficients of the objective function greater than

0 by fixing factors H and $\{W_1, W_2, L, W_r\} / W_i$. In contrast to other methods, this kind of NMF has a faster convergence rate. However, the changes caused by ε affect the attributes of the results of decomposition. As a negative result, sparsity cannot be achieved, which impedes the application of the NMF algorithm in the analysis of hyperspectral remote sensing. Given the fact that only several kinds of endmembers often appear in the pixels of a specific hyperspectral image, the previously mentioned NMF algorithm, which remains non-negative with certain constraints, could be modified to achieve an INMF algorithm.

In the non-negative model, Formula (4) can be cast into Formulas (5) and (6) by fixing $(H, \{W_1, W_2, L, W_r\} / W_i)$ and $(W, \{H_1, H_2, L, H_r\} / H_j)$, respectively. Formulas (5) and (6) are expressed as follows:

$$\min_{W_i \geq 0} f(W_i) = \left\| V - \sum_{k=1, k \neq i}^r W_k H_k - W_i H_i \right\|_F^2, \quad (5)$$

$$\min_{H_j \geq 0} f(H_j) = \left\| V - \sum_{k=1, k \neq j}^r W_k H_k - W_j H_j \right\|_F^2. \quad (6)$$

The terms that contain only W_{ci} , which is expanded from Formula (5), are independent of variables as $\{W_{1i}, W_{2i}, L, W_{mi}\} / W_{ci}$. Thus, the objective function that consists of only W_{ci} could be simplified as follows:

$$\begin{aligned} f(W_{ci}) &= \left\| V_c - \sum_{k=1, k \neq i}^r W_{ck} H_k - W_{ci} H_i \right\|_2^2 \\ &= \sum_{j=1}^n (V_{cj} - \sum_{k=1, k \neq i}^r W_{ck} H_{kj} - W_{ci} H_{ij})^2 \\ &= (HH^T)_{ii} W_{ci}^2 - 2\eta_{ci} W_{ci} + \gamma \end{aligned} \quad (7)$$

where $\eta_{ci} = (VH^T)_{ci} - \sum_{k=1, k \neq i}^r W_{ck} (HH^T)_{ki}$ and $\gamma = \sum_{j=1}^n (V_{cj} -$

$\sum_{k=1, k \neq i}^r W_{ck} H_{kj})^2$ are constants that are entirely unrelated to W_{ci} . Formula (7) shows that, for $(HH^T)_{ii} \neq 0$, if $W_{ci} \geq 0$, then $W_{ci} = \max\{\eta_{ci} / (HH^T)_{ii}, 0\}$. For $(HH^T)_{ii} = 0$, if $\|H_i\| > 0$, then $W_{ci} = \max\{\eta_{ci} / (HH^T)_{ii}, 0\}$; else, if $\|H_i\| = 0$, then W_{ci} could be valued as a random number from 0 to 1. Thus, functions from $f(W_{c1})$ to $f(W_{cr})$ are uncorrelated. Thus, we derive the following expression when the minimum of $f(W_i)$ is reached:

$$W_i = \begin{cases} \max\{\eta_i / (HH^T)_{ii}, 0\}; & \|H_i\| > 0 \\ rand(m, 1); & \|H_i\| = 0 \end{cases}, \quad (8)$$

where $rand(m, 1)$ denotes a m -dimensional random column vector that takes a value between 0 and 1, and

$$\eta_i = (\eta_{1i}, \eta_{2i}, L, \eta_{mi})^T = (VH^T)_i - \sum_{k=1, k \neq i}^r W_k (HH^T)_{ki}. \quad (9)$$

The following conclusions can be drawn based on the previously presented analysis.

Formula (8) holds for any given $i \in \{1, 2, L, r\}$ by fixing H and $\{W_1, W_2, L, W_r\} / W_i$ when $f(W_i)$ is minimized to the limit value. Similarly, we derive the following expression fixing W and $\{H_1, H_2, L, H_r\} / H_j$ when the minimum of $f(H_j)$ is reached for any $j \in \{1, 2, L, r\}$:

$$H_j = \begin{cases} \max\{\xi_j / (W^T W)_{jj}, 0\}; & \|W_j\| > 0 \\ 0_{1 \times n}; & \|W_j\| = 0 \end{cases}, \quad (10)$$

where $\xi_j = (W^T V)_j - \sum_{k=1, k \neq j}^r (W^T W)_{jk} H_k$.

The procedure of the proposed INMF algorithm is summarized in Table 1.

Table 1. Procedure of the INMF algorithm.

Step 1) Given V and r , initialize $W(0) \geq 0, H(0) \geq 0$ randomly, let $k = 0$.
Step 2) Set $D(k) = H(k) H^T(k), Q(k) = V H^T(k), E(k) = D(k)$. For $i = 1, 2, \dots, r, E_{ii}(k) = 0$, if $D_{ii}(k) > 0$, then $\hat{W} = (W_{\cdot 1}(k+1) \dots W_{\cdot i-1}(k+1) \dots W_{\cdot i}(k) \dots W_{\cdot r}(k)), W_{\cdot i}(k+1) = \max\{(Q_{\cdot i}(k) - \hat{W} E_{\cdot i}(k)) / D_{ii}(k), 0\}$, else $W_{\cdot i}(k+1) = rand(m, 1)$.
Step 3) Set $C(k) = W^T(k+1)W(k+1), R(k) = W^T(k+1)V, F(k) = C(k)$. For $j = 1, 2, \dots, r, F_{jj}(k) = 0$, if $C_{jj}(k) > 0$, then $\hat{H} = (H_{\cdot 1}(k+1) \dots H_{\cdot j-1}(k+1) \dots H_{\cdot j}(k) \dots H_{\cdot r}(k)), H_{\cdot j}(k+1) = \max\{(R_{\cdot j}(k) - F_{\cdot j}(k)) \hat{H} / C_{jj}(k), 0\}$, else $H_{\cdot j}(k+1) = 0_{1 \times n}$.
Step 4) $k = k + 1$, repeat steps 2 and 3 while k is less than the stopping condition.

As discussed in this section, the optimization of the NMF can be transformed into a series of independent subproblems in vectorized form to obtain the minimum of the objective function. Furthermore, the function scale reduces rapidly because of the fact that the subproblems are independent and the terms that are unrelated with these subproblems are fixed. Thus, the proposed INMF algorithm achieves fast convergence under non-negative constraint.

3.3 Fusion model based on INMF

The original NMF problem of Equation (4) may suffer from the impossibility of acquiring the global optimal solution because of the nonconvexity of the cost function with respect to W and H . Some a priori should be imposed on the cost function according to the actual condition to make the solution more reasonable. Hence, sparseness is the first constraint that should be added in the case of HSRI unmixing.

When the cost function (Equation (4)) is utilized to deal with hyperspectral image fusion, the mathematical notations can be defined as follows. V is the source HSRI with L bands and N pixels and column vector $\{V_j\}_{j=1}^N \in \mathbb{R}^L$ denotes any given pixel with L color channels. Endmember matrix W is composed of a certain pure material spectrum, such as forest, soil, and water with each column standing for a specific spectrum. H is the abundance matrix written as $\{H_j\}_{j=1}^N \in \mathbb{R}^P$, wherein each column denotes the fractional abundance corresponding to the specific endmember. P represents the number of endmembers. The abundance of any endmember is not always equally distributed. Thus, only

a small part of endmembers is necessary when reconstructing one pixel of V . Consequently, the L_1 -norm-based constraint should be imposed on factor H to make it sparse, as shown in Formula (11):

$$\min_{W,H} f(W,H) = \frac{1}{2} \left\{ \|V - WH\|_F^2 + \beta \sum_{j=1}^N \|H_j\|_1 \right\}, \text{ s.t. } W, H \geq 0, \quad (11)$$

where $\|H_j\|_1 = \sum_{i=1}^P |H_{ij}|^2$ is the L_1 -norm vector and β is a user-defined parameter that balances the trade-off between the accuracy of the reconstruction and the sparseness.

3.4 Fusion method

According to descriptions of hyperspectral image fusion presented in Sections 3.1 and 3.2, the fusion method should meet at least two requirements to obtain an integrated HSGRI with little spectral distortion. Aside from the sharpening information, HGRI or high-frequency data must be injected into the low spatial resolution HSRI to ensure that HSGRI has more geometric information. Explicit measures must be implemented to ensure that any single pixel of HSGRI approximates that of HGRI as close as possible in the spectrum aspect.

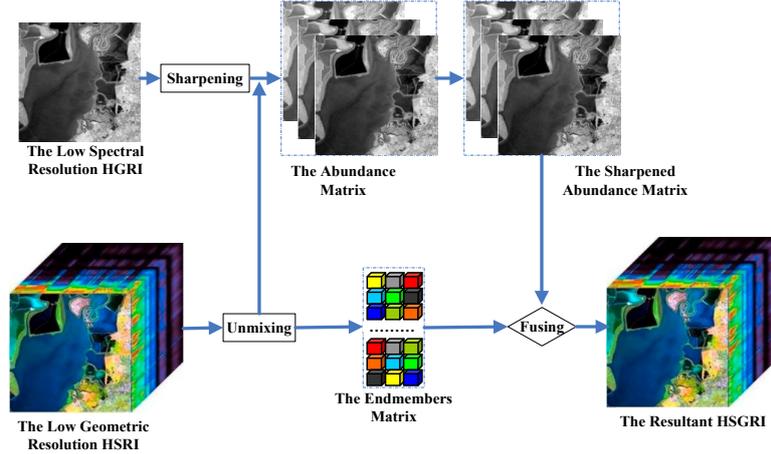


Fig. 1. Hyperspectral image fusion process.

The original low geometric resolution HSRI can be decomposed into the endmember matrix W and the corresponding fractional abundance matrix H . H clearly contains the low geometric resolution information from V . Therefore, high geometric resolution HSRI can be reasonably generated by means of sharpening H with the high geometric resolution information from HGRI. Fig. 1 illustrates the process of hyperspectral image fusion, which can be regarded as the optimization of Formula (11). Finally, the ultimate fused HSGRI is produced, denoted as V_f and expressed as follows:

$$V_f = W(\alpha H + (1 - \alpha) \text{HGSIs}), \quad (12)$$

where $\text{HGRIs} \in \mathbb{R}^{P \times N}$, and the notation can be reformulated row by row as $\{\text{HSRIs}_j\}_{j=1}^P \in \mathbb{R}^N$, which denotes the duplication of the HGRI. Parameter α is used to control the degree to which the abundance is preserved and the HSRI is sharpened.

The spectral preservation of the original hyperspectral image should also be considered. Although the spatial information from the HGRI is injected into the HSGRI in this model, it does not mean that the spectral quality of the HSGRI is totally identical to that of the HSRI. Generally, the spectral information of the HSGRI differs from that of the HSRI in many cases. Thus, spectral distortion occurs. Consequently, we need to use spectral preservation terms to constrain the discrepancy between the spectra of HSRI and HSGRI to reduce the degree of distortion.

As one of the measuring terms in remote sensing, spectral angle (SA) is widely used to assess the spectral

similarity between HSRI and HSGRI, which is defined as follows:

$$SA_{UV} = \arccos \left(\frac{\sum_i^k u_i v_i}{\sqrt{\sum_i^k u_i^2} \sqrt{\sum_i^k v_i^2}} \right), \quad (13)$$

where k denotes the number of pixels and U_i and V_i represent the spectral vectors of image U and V at pixel k , respectively. The greater the value of SA_{UV} , the smaller the similarity between U and V will be. However, SA is not always sensitive in discriminating image differences, particularly when these images are highly similar. SA is optimized to improve its sensitivity. This term is called enhanced SA (ESA) in this context, which is defined as follows:

$$ESA_{UV} = \arccos \left(2 \frac{\sum_i^k u_i v_i}{\sum_i^k u_i^2 + \sum_i^k v_i^2} \right). \quad (14)$$

As shown in Formulas (13) and (14), aside from the different values of denominators and numerators, ESA is similar to SA in mathematical form. Index ESA , which deals with isolated pixels only and does not consider the relationship between neighboring pixels, is important for finding differences in structures and textures. Compared with SA , ESA is more sensitive to the differences between two images as the ESA equals 0 only when two images are identical, whereas the SA can be 0 when two images are similar. For example, when $U = \{u_1, u_2, \dots, u_i\}$, $V = \{v_1, v_2, \dots, v_i\}$, and $u_i \neq v_i$, the value of SA is 1, but the value

of *ESA* is less than 1. Therefore, the *ESA* is more informative than the *SA* in measuring the degree of similarity quantitatively.

For Formula (14), u_i can be replaced by v_i^f . We then derive Formula (15), where v_i^f and v_i refer to the images after and before fusion, respectively. In this study, they denote the HSGRI and HSRI, respectively.

$$ESA = \arccos \left(2 \sum_i^k v_i^f v_i / \left(\sum_i^k v_i^f{}^2 + \sum_i^k v_i^2 \right) \right). \quad (15)$$

Given that Formula (15) is complex and difficult to implement, the anti-trigonometric function is removed. For simplicity, Formula (15) is transformed into the form of a function, as expressed in Formula (16):

$$ESA(V_f, V) = \sum_i^k (2 \langle v_i^f, v_i \rangle / (\|v_i^f\| + \|v_i\|)). \quad (16)$$

Better spectral preservation results in smaller *ESA*. If V_f and V are absolutely the same, then spectral distortion will not occur.

The function for spectral preservation, expressed in Formula (16), is added to the L_1 -norm-enforced Formula (11). Thus, the objective function is imposed by the sparse and spectral constraints, as shown in Formula (17):

$$\min_{W, H} C(W, H) = \frac{1}{2} \left\{ \|V - WH\|_F^2 + \beta \sum_{j=1}^N \|H_j\|_1^2 + \lambda ESA(V_f, V) \right\}, \quad (17)$$

s.t. $W, H \geq 0$

where γ is a constant used to balance the spectral and spatial qualities of the fused HSGRI. Thus, the proposed method not only improves the spatial resolution of HSGRI by extracting spatial information from HGRI, but also preserves the spectral features of the HSRI relying on the spectral constraint function $ESA(V_f, V)$.

3.5 Fusion implementation

The objective function [Equation (17)] has three addends, with the latter two constrained. In fact, several existing studies add ‘‘penalty terms’’ to the NMF objective function to improve performance and derive the following normal form:

$$\min_{W, H} = \left\{ \|V - WH\|_F^2 + \lambda \|W\|_F^2 + \beta \sum_{j=1}^N \|H_j\|_1^2 \right\}, \quad \text{s.t. } W, H \geq 0. \quad (18)$$

Concretely speaking, an alternative algorithm is considered a general method to solve Formula (18). However, the matrix size of the original problem is amplified, which overburdens calculation. In the proposed INMF method, the objective function for W_{ci} is

$$f_1(W_{ci}) = [(HH^T)_{ii} + \lambda] W_{ci}^2, \quad \text{where } \eta_{ci} = (VH^T)_{ci} - \sum_{k=1, k \neq i}^r W_{ck} (HH^T)_{ki},$$

γ_1 is a constant with respect to W_{ci} , and W_{ci} can be calculated as $W_{ci} = \max\{\eta_{ci} / [(HH^T)_{ii} + \lambda], 0\}$. For the objective function of H_{jf} , $f_1(H_{jf}) = [(W^T W)_{jj} + \beta] H_{jf}^2 - 2\xi_{jf} H_{jf} + \gamma_2$, where $\xi_{jf} = (W^T V)_{jf} - \sum_{k=1, k \neq j}^r [(W^T W)_{jk} + \beta] H_{kf}$,

γ_2 is a constant with respect to H_{jf} , and H_{jf} can be calculated as $H_{jf} = \max\{\xi_{jf} / [(W^T W)_{jj} + \beta], 0\}$. Accordingly, the updating rules for the model are defined as follows:

$$W_i = \max\{\eta_i / [(HH^T)_{ii} + \lambda], 0\}$$

$$H_j = \max\{\xi_j / [(W^T W)_{jj} + \beta], 0\}. \quad (19)$$

Formulas (17) and (18) have a similar mathematical form. Therefore, the ultimate fusion model can be achieved by replacing the variables in Formula (18) and updating rules in Formula (19) with the corresponding counterparts in Formula (17).

4 Result analysis and discussion

Two groups of hyperspectral images, i.e., synthetic and real data, are selected to conduct the experiments under the platform of MATLAB to verify the effect of the proposed algorithm. In this experiment, the performance of the proposed INMF algorithm is compared with that of high-pass filtering (HPF), discrete wavelet transform (DWT), intensity hue saturation (IHS), Ehler, and classical NMF.

The quality of the fusion image can be evaluated visually. However, the result of this method may differ because subjective evaluation depends on the physical condition, psychological status, and even mental states of the observers. Therefore, the objective criteria for HSGRI, original HSRI, and reference HSGRI should be considered in quantitative evaluation. The following criteria are used in this study: such as information entropy (*IE*), root mean square error (*RMSE*), expanded spectral angle mapper, correlation coefficient (*CC*), average gradient (*AG*), universal image quality index (*UIQI* or *Q-average*), spectral information divergence (*SID*), and relative dimensionless global error in synthesis (*ERGAS*).

4.1 Evaluation criteria

Information Entropy (*IE*)

The entropy of an image is defined as follows:

$$E = - \sum_{i=1}^N P_i \log P_i, \quad (20)$$

where i is the gray value of an image from 1 to N and P_i denotes the probability of a specific i . *IE* is one of the most important indices for evaluation, whose value directly reflects the amount of information in the image. The higher the value of *IE* is, the higher the spatial resolution in a fused HSGRI.

Average Gradient (*AG*)

AG reflects the contrast and intelligibility of an image and is defined as follows:

$$AG = \frac{1}{(M-1)(N-1)} \sum_{i=1}^M \sum_{j=1}^N \sqrt{\frac{(\nabla x)^2 + (\nabla y)^2}{2}}, \quad (21)$$

where M and N represent the rows and columns of the image, respectively, and ∇x and ∇y are the gradients along the x - and y -directions, respectively. Generally, the larger the value of *AG* is, the higher the definition of the image.

Correlation Coefficient (*CC*)

The *CC* between HSGRI and original HSRI measures the similarity of the spectra and is defined as follows:

$$CC = \frac{\sum_{i=1}^M \sum_{j=1}^N [H(i,j) - \bar{H}][H_f(i,j) - \bar{H}_f]}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N [H(i,j) - \bar{H}]^2 \sum_{i=1}^M \sum_{j=1}^N [H_f(i,j) - \bar{H}_f]^2}}, \quad (22)$$

where $H(i,j)$ and $H_f(i,j)$ represent the gray value of a pixel at (i,j) in the HSRI and HSGRI with the size of $M \times N$, respectively, and \bar{H} and \bar{H}_f denote the mean gray value of all pixels with respect to HSRI and HSGRI, respectively. The higher the value of *CC* is, the better the spectral quality of the fused HSGRI. *CC* close to 1 indicates that the two images are highly correlated.

Universal Image Quality Index (*Q*-average)

The metric *Q*-average (abbreviated as *Q*) is used to assess the quality of image sharpening and is defined as follows:

$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x}^2 + \bar{y}^2)} \cdot \frac{2\sigma_x \sigma_y}{(\sigma_x^2 + \sigma_y^2)}, \quad (23)$$

where \bar{x} , \bar{y} and σ_x , σ_y are the mean values and standard deviations of the images x and y , respectively. The index of *Q* reflects the combination of three factors, namely, loss of correlation, luminance distortion, and contrast distortion. The perfect value is 1.

Average Expanded Spectral Angle (*AESA*)

ESA is a particularly informative metric in terms of measuring how close the pixel values of the two images are. *ESA* is defined in Formula (14), where k denotes the number of pixels and U_i and V_i represent the spectral vectors of images U and V at pixel k , respectively. *AESA* is adopted as a holistic index for measuring the spectral differences between original HSRI and fused HSGRI. The higher the value of *AESA* is, the lesser the similarity of the two images will be.

Spectral Information Divergence (*SID*)

The index of *SID* can be used to describe the spectral correlation between two images. Given that the vectors $x = (x_1, x_2, \dots, x_L)$ and $y = (y_1, y_2, \dots, y_L)$ are pixels taken from the fused HSGRI and original HSRI, we derive the following expression:

$$SID(x,y) = D(x||y) + D(y||x), \quad (24)$$

where $D(x||y)$, $D(y||x)$ indicate the relative entropy of y and x with respect to x and y , which are defined as follows:

$$D(y||x) = \sum_{l=1}^L p_l \log\left(\frac{p_l}{q_l}\right), \quad D(x||y) = \sum_{l=1}^L q_l \log\left(\frac{q_l}{p_l}\right), \quad (25)$$

where:

$$p_l = \frac{x_l}{\sum_{i=1}^L x_i}, \quad q_l = \frac{y_l}{\sum_{i=1}^L y_i}. \quad (26)$$

Root Mean Square Error (*RMSE*)

The *RMSE* is commonly used to measure the error between two images by directly calculating the changes in pixel values and is defined as follows:

$$RMSE = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (H(x,y) - H_f(x,y))^2}, \quad (27)$$

where $H(x,y)$ and $H_f(x,y)$ denote the pixel value of the original HSRI and fused HSGRI, respectively, and M and N represent the rows and columns of the image, respectively. The fused HSGRI is closer to the original HSRI when the *RMSE* is small and ideally expected to be 0.

Relative Dimensionless Global Error in Synthesis (*ERGAS*)

The *ERGAS* reflects the overall quality of the HSGRI, indicates the difference between original HSRI and fused HSGRI and is defined as follows:

$$ERGAS = 100 \frac{h}{l} \sqrt{\frac{1}{L} \sum_{k=1}^L \left[\frac{RMSE_k}{\bar{X}_k} \right]^2}, \quad (28)$$

where h and l denote the resolutions of HGRI and HSRI, respectively, L is the number of bands, and $RMSE_k$ and \bar{X}_k are the *RMSE* and mean values of the k^{th} band, respectively. A small *ERGAS* indicates that a small spectral distortion occurs and the algorithm is more helpful in spectral preservation.

4.2 Experimental results with synthetic datasets

Hyperspectral data over Texas are obtained from the U.S. Army Geospatial Center [31]; these data are composed of 210 bands with the spectral resolution of 10 nm and the size of 307×307 collected by the hyperspectral digital imagery collection sensor. The noisy bands of 105-109, 139-151, and 207-210 are removed, and 80 bands from the rest of the data are randomly selected to constitute an image cube. From this cube, band 215 is selected and simulated as the panchromatic image and is taken as the HGRI in that this band has a higher spatial resolution and definition compared with the other bands. For the acquisition of simulated HSRI, the original hyperspectral data are low-pass filtered, downsampled by a factor of 4, and resized to its original size. Panchromatic and hyperspectral images are obtained through this method. Given that both images are taken from the same data source, they are naturally registered.

As shown in Figs. 2(a) and 2(b), the holistic representation of the hyperspectral image is fuzzy and out of focus, whereas that of the panchromatic image is distinct. Hyperspectral image and all of the fused images displayed in Fig. 2 are composed of bands 50, 133, and 183 for the red, green, and blue colors respectively. Five comparative approaches plus the proposed INMF are applied to verify the fusion performance. Figs. 2(c) to 2(h) show the fused results.

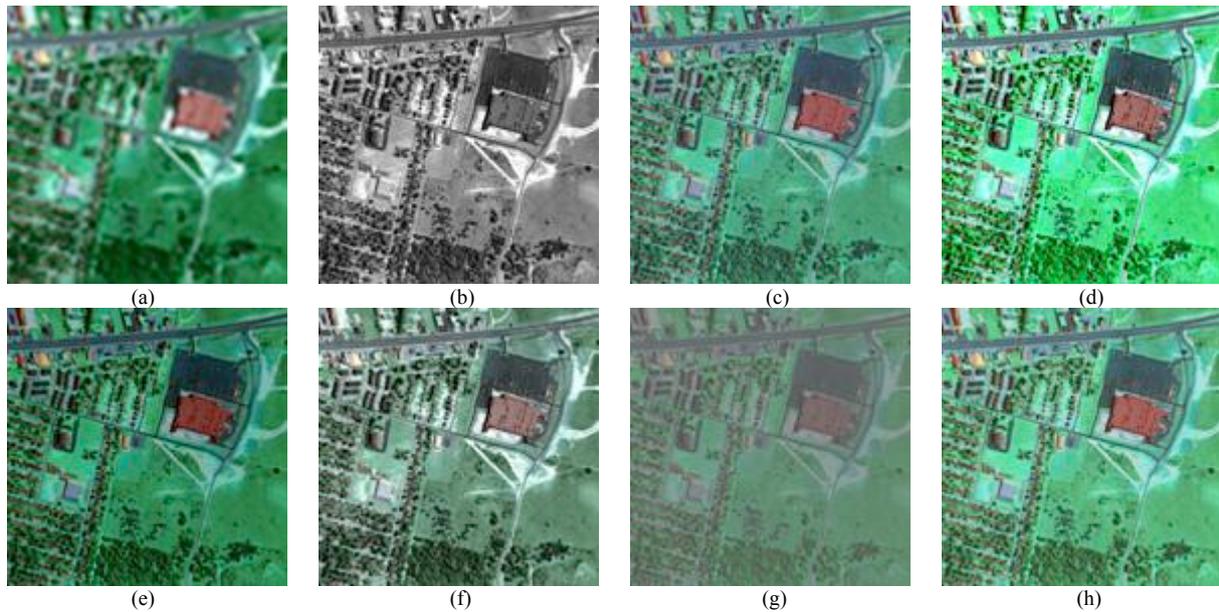


Fig. 2. Comparison of the fusion methods of synthetic images: (a) hyperspectral image, (b) panchromatic image, and (c) to (h) fused images obtained using the HPF, DWT, IHS, Ehler, classical NMF, and proposed INMF.

The fusion methods integrate the characteristics of both source images. As far as the spatial resolution is concerned, all fused images are spatially enhanced, i.e., the houses in the upper left, the cars parking on the ground, and the roads on the top and right of the images can be identified. As for the spectral resolution, Figs. 2(c), 2(f), and 2(g) show images with a low contrast. Thus, Figs. 2(c) and 2(g) are dim and Fig. 2(f) looks washed out. These instances can all be

regarded as spectral (color) distortion. Figs. 2(d), 2(e), and 2(h) have better spectral preservation effects, as these three images are similar to the original hyperspectral image in the overall hue presentation. However, Fig. 2(d) has a high intensity, which makes it unusual visually. The results of the objective criteria are listed in Table 2 to obtain a quantitative statistic.

Table 2. Evaluation results of the simulated data.

	HPF	DWT	IHS	Ehler	Classical NMF	INMF	Optimum value
<i>IE</i>	7.3169	7.7111	7.2914	7.6217	6.5310	7.7785	7.7785
<i>AG</i>	0.0628	0.1083	0.0632	0.0742	0.0376	0.0942	0.1083
<i>CC</i>	0.8043	0.8041	0.8025	0.8078	0.8086	0.8115	1
<i>Q</i>	0.8246	0.8546	0.8756	0.8993	0.8921	0.9075	1
<i>AESA</i>	0.3347	0.4691	0.3822	0.3586	0.3987	0.3455	0
<i>SID</i>	0.0618	0.0725	0.2253	0.0293	0.0345	0.0128	0
<i>RMSE</i>	1.2326	0.9871	4.6179	6.4126	0.7158	0.5643	0
<i>ERGAS</i>	13.5284	15.6512	18.0654	31.5871	22.4512	9.7324	0

Table 2 summarizes the results of the assessment for the eight criteria. The last column shows the reference values and the best value in each criterion rendered in boldface. Notably, the proposed INMF achieves the best performance in six of eight metrics, and the classical NMF has the worst performance in *IE*, *AG*, and *AESA*, which is in accordance with its visual effect. For *CC* and *SID*, the proposed method achieves the best value, which is slightly larger than that of HPF in *AESA*, indicating that the INMF is superior to the other methods in spectral preservation. For spatial rendering, the proposed method has the best *Q* index and is at the second place in terms of *AG* when compared with that of DWT. In the *RMSE* and *ERGAS* indices, the proposed INMF also has the smallest value, but these two values are not close enough to 0 because the original hyperspectral image is downsampled and random errors are introduced. In columns 4 and 5, IHS is the worst in *CC* and *SID*, whereas Ehler is the worst in *RMSE* and *ERGAS*, which are closely related to the subtle spectral distortion shown in Fig. 2(e) and the unnatural artifacts shown in Fig. 2(f), respectively. In brief, the proposed INMF exhibits a better balance between spectral preservation and spatial enhancement in synthetic image fusion.

4.3 Experimental results with real datasets

Aside from synthetic image fusion, experiments on real images are also conducted. The images used in this part of the experiments show a view of airborne flight over the beach region of southern Greece with pixel response in 189 bands from the 0.4 μm to the 2.4 μm region of the visible and infrared spectra downloaded from the Leica Geosystems Geographic Imaging website [32], where bands from 0.8 μm to 1.3 μm have been removed. The original hyperspectral image is 296×296 with the spatial resolution of 100 m, whereas the panchromatic image is 512×512 with the spatial resolution of 30 m. All fused and hyperspectral images shown in Fig. 3 are made using bands 15, 27, and 48 for the red, green, and blue colors respectively.

Similarly, the six fusion methods employed in synthetic data are also used to evaluate the performances of these algorithms. As shown in Figs. 3(a) and 3(b), the panchromatic image shows the well-defined coastal road, rugged hill ridges, clear outline of the houses, and a circular construction located at the middle bottom, whereas the hyperspectral image, with one third of the picture filled with seawater in dark blue, is foggy and fails to illustrate the previously mentioned details. These fused images produced

by these six methods look more distinct than the original hyperspectral image, and the previously mentioned objects can be identified. All fused images have abundant spectral information similar to the hyperspectral image does. Notably, the IHS has a different hue presentation and turns the dark blue water shown in Fig. 3(a) into dark green, as shown in

Fig. 3(e), when compared with the other methods, which can be regarded as spectral distortion. As shown in Fig. 3(g), the houses opposite to the beach, the triangle area between the two highways, and the hills on middle left are all blurred, which makes the image visually unique.

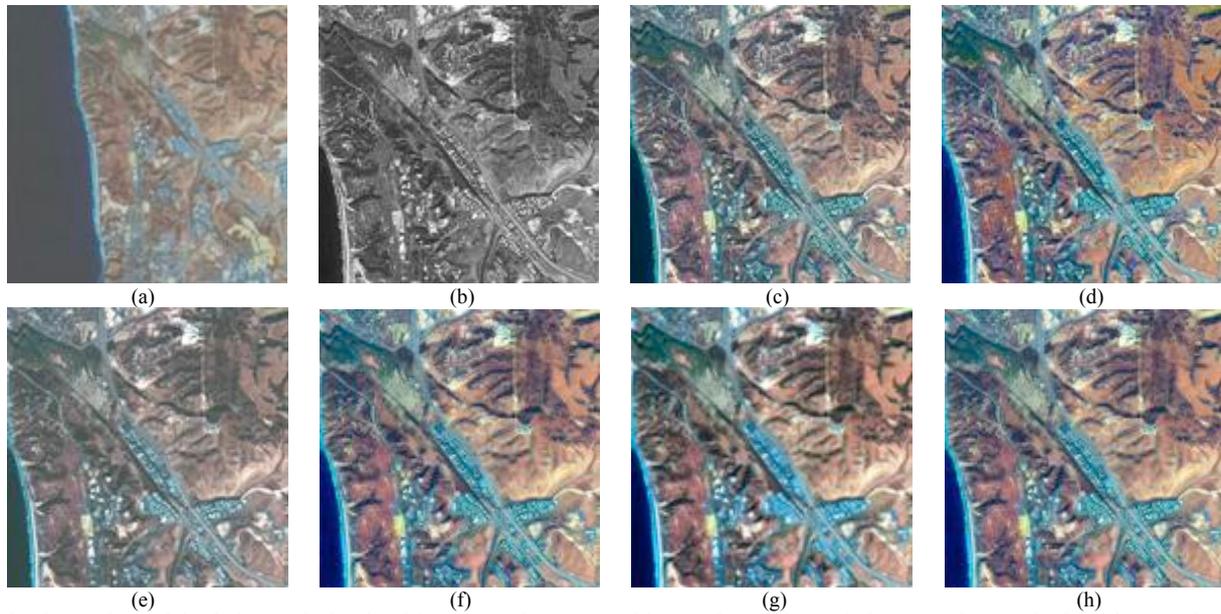


Fig. 3. Comparison of the fusion methods of real images: (a) hyperspectral image, (b) panchromatic image, and (c) to (h) fused images obtained using HPF, DWT, IHS, Ehler, classical NMF, and proposed INMF.

Table 3. Evaluation results of real data.

	HPF	DWT	IHS	Ehler	Classical NMF	INMF	Optimum value
<i>IE</i>	8.3412	8.5086	8.2424	8.5569	8.2671	8.5681	8.5681
<i>AG</i>	0.0952	0.1041	0.1015	0.0898	0.0923	0.1038	0.1041
<i>CC</i>	0.8138	0.8105	0.8075	0.8125	0.8098	0.8956	1
<i>Q</i>	0.9478	0.9445	0.9441	0.9373	0.9144	0.9847	1
<i>AESA</i>	0.3915	0.3607	0.4012	0.3998	0.3976	0.3072	0
<i>SID</i>	0.0654	0.0277	0.0428	0.0124	0.0368	0.0115	0

Quantitative estimations are made and the results are listed in Table 3 to further verify the visual effect. The criteria for *RMSE* and *ERGAS* do not appear because the standard reference image is not available. The proposed algorithm achieves the first place in five criteria except the index *AG*. In terms of spectral preservation, the proposed INMF exceeds that of HPF by 10.05% in *CC*, and as for *AESA* and *SID*, INMF is superior to that of DWT and Ehler by 17.41% and 7.82%, respectively. In spatial resolution enhancement, the proposed algorithm takes the first place in *Q*, and surpasses that of HPF by 3.89%. HPF, DWT, Ehler and the proposed algorithm have little difference in spectral preservation and spatial improvement, but IHS holds the final place as far as the amount of information and the ability to preserve spectrum are concerned. The classical NMF is the last one in *AG* and *Q*, which implies that it has the poorest performance in spatial presentation.

4.4 Numerical experiment and analysis

The runtime of the proposed algorithm is compared with that of wavelet and classical NMF under the platform of Athlon 2.61 MHz, 2 GB RAM to analyze the advantage of application for hyperspectral data. The initial hyperspectral and panchromatic data are generated by cropping square matrices with the size of 64×64 similar to those in Figs. 3(a) and 3(b). First, the proposed INMF algorithm is implemented until it ends, and the time elapsed is recorded.

Then, the classical NMF and wavelet are implemented under the same conditions, and the time elapsed is recorded. Subsequently, the number of pixels for the row and column of the initial 64×64 matrix are doubled gradually until the data volume reaches approximately 1,000 MB ($32,768 \times 32,768$). As shown in Fig. 4, the Y-axis represents the scale of the time consumed and the X-axis denotes the data size. Fig. 4 shows that the three algorithms are not time consuming in the early stage for data of small scale. However, when the number of pixels reaches 1,024, the differences of the runtime of these three methods can be recognized. The proposed algorithm takes the least time and the classical NMF algorithm needs the most time. When the number of pixels per line exceeds approximately 22,000 (the scale in the X-axis at the intersection), the time consumption of the proposed algorithm surpasses that of wavelet and tends to increase rapidly with the classical NMF. In other words, the two NMF algorithms converge fast under the condition of small data. However, when the size reaches the threshold of 480 MB at the intersection in this context, the convergence rate becomes slow because of the multiplication of the image size. In summary, two algorithms based on the NMF may converge to the minimum in 10 s when the data size is less than 16 MB ($4,096 \times 4,096$), but are not fast enough for a large batch, massive data size, and real-time operation unless the hardware speed is considerably improved.

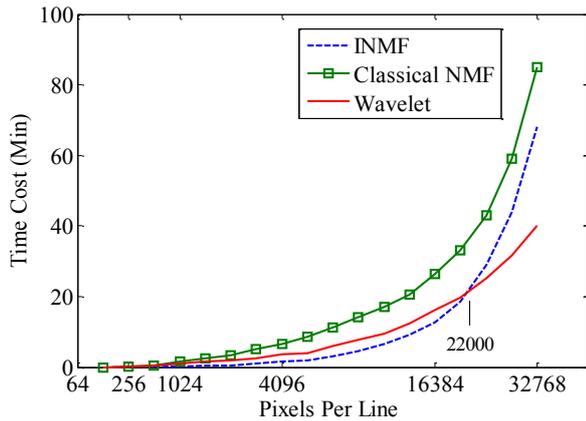


Fig. 4. Runtime comparison of the three fusion methods.

5. Conclusions

Spectral and sparse constraints were imposed on the cost function to preserve the spectral features and enhance the spatial resolution in hyperspectral image fusion. A model based on the INMF algorithm was developed to optimize the efficiency of the proposed method. Databases of the U.S. Army Geospatial Center and Leica Geosystems Geographic Imaging were used to evaluate the validity of the proposed method. The following conclusions could be drawn

(1) The spatial resolution of the fused image is intensified by extracting details from the high spatial resolution image. Techniques of sparseness and sharpness control enable the proposed method to improve the spatial resolution of the fused image.

(2) The function of spectral constraint, which quantifies the spectral similarity between original hyperspectral and fused images, is capable of preserving the spectral features of the original hyperspectral image.

(3) The process of NMF optimization could be converted into several subproblems on the premise that non-negativity is ensured. The proposed INMF method achieves fast matrix decomposition under the condition of sparse and spectral constraints. When the image size is less than 480 MB, the proposed method could be applied as an effective fusion scheme for hyperspectral images.

Theoretical research and experimental comparison are combined to propose a fast INMF algorithm. The proposed model based on spectral and sparse constraints with clear physical meaning and ease of implementation is highly applicable to the practice of hyperspectral image fusion. Compared with the traditional fusion methods, the proposed method produces equivalent or better images and consumes less time. However, the manner of random initialization, the diversity of endmembers, and the vast data scale have adverse effects on the model. Therefore, the algorithm must be optimized further to enable large-scale and real-time application to networks.

Acknowledgements

This work was supported by the Project of the Sichuan Provincial Department of Education (Grant No. 152A0148).

This is an Open Access article distributed under the terms of the Creative Commons Attribution License



References

- Lanaras, C., Baltsavias, E., Schindler, K., "Hyperspectral super-resolution by coupled spectral unmixing". In: *Proceedings of the 15th IEEE International Conference on Computer Vision*, Santiago, Chile: IEEE, 2015, pp. 3586-3594.
- Altmann, Y., Halimi, A., Dobigeon, N., et al., "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery". *IEEE Transactions on Image Processing*, 21(6), 2012, pp. 3017-3025.
- Gonzalez, C., Resano, J., Plaza, A., et al., "FPGA implementation of abundance estimation for spectral unmixing of hyperspectral data using the image space reconstruction algorithm". *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(1), 2012, pp. 248-261.
- Zare, A., Ho, K. C., "Endmember variability in hyperspectral analysis: addressing spectral variability during spectral unmixing". *IEEE Signal Processing Magazine*, 31(1), 2014, pp. 95-104.
- Wang, J., Lai, S.Y., Li, M. D., "Improved image fusion method based on NSCT and accelerated NMF". *Sensors*, 12(5), 2012, pp. 5872-5887.
- Dalponte, M., Bruzzone, L., Gianelle, D., "Tree species classification in the southern Alps based on the fusion of very high geometrical resolution multispectral / hyperspectral images and LiDAR data". *Remote Sensing of Environment*, 123(1), 2012, pp. 258-270.
- Bioucas-Dias, J. M., Plaza, A., Camps-Valls, G., et al., "Hyperspectral remote sensing data analysis and future challenges". *IEEE Geoscience & Remote Sensing Magazine*, 1(2), 2013, pp. 6-36.
- Grohnfeldt, C., Zhu, X., X., Bamler, R., et al., "Jointly sparse fusion of hyperspectral and multispectral imagery". In: *Proceedings of the International Geoscience and Remote Sensing Symposium*, Melbourne, Australia: IEEE, 2013, pp. 4090-4093.
- Kim, Y., "Generalized IHS-based satellite imagery fusion using spectral response functions". *Etri Journal*, 33(4), 2011, pp. 497-505.
- Mandhare, R. A., Upadhyay, P., Gupta, S., "Pixel-level image fusion using brovey transform and wavelet transform". *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2(6), 2013, pp. 2690-2695.
- Licciardi, G. A., Khan, M. M., Chanussot, J., et al., "Fusion of hyperspectral and panchromatic images using multiresolution analysis and nonlinear PCA band reduction". *EURASIP Journal on Advances in Signal Processing*, 2012(1), 2012, pp. 1783-1789.
- Chisense, C., Engels, J., Hahn, M., et al., "Pansharpening of hyperspectral images in urban areas". In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Melbourne, Australia: 2012, pp. 387-392.
- Mansour, N., Shadrokh, S., Shahram, S., "Multi-focus image fusion using dictionary-based sparse representation". *Information Fusion*, 25(9), 2015, pp. 72-84.
- Wei, Q., Dobigeon, N., Tourneret, J., "Bayesian fusion of multispectral and hyperspectral images using a block coordinate descent method". In: *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Tokyo, Japan: IEEE, 2015, pp. 1254-1259.
- Shoyaib, M., Abdullah, M., Chae, O., "A skin detection approach based on the Dempster-Shafer theory of evidence". *International Journal of Approximate Reasoning*, 53(4), 2012, pp. 636-659.
- Das, S., Kundu, M. K., "NSCT based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency". *Medical & Biological Engineering & Computing*, 50(10), 2012, pp. 1105-1114.
- Wang, J., Wang, H. J., Li, Y. H., et al., "Image fusion and evaluation of geological disaster based on remote sensing". *International Journal of Online Engineering*, 10(4), 2014, pp. 28-34.
- Tian, J., Chen, L., "Adaptive multi-focus image fusion using a wavelet based statistical sharpness measure". *Signal Processing*, 92(9), 2012, pp. 2137-2146.

19. Choi, Y., Sharifahmadian, E., Latifi, S., "Performance analysis of contourlet based hyperspectral image fusion methods". *International Journal on Information Theory*, 2(1), 2013, pp. 1-14.
20. Chen, G., Qian, S. E., Ardouin, J. P., et al., "Super resolution of hyperspectral imagery using complex ridgelet transform". *International Journal of Wavelets, Multi-resolution and Information Processing*, 10(3), 2012, pp. 1-22.
21. McLaughlin, M. J., Grieggs, S., Ezekiel, S., et al., "Bandelet Denoising in Image Processing". In: *Proceedings of the 2015 IEEE National Aerospace and Electronics Conference*, Ohio, USA: IEEE, 2015, pp. 35-40.
22. Halimi, A., Altmann, Y., Dobigeon, N., et al., "Nonlinear unmixing of hyperspectral images using a generalized bilinear model". *IEEE Transactions on Geoscience & Remote Sensing*, 49(11), 2014, pp. 4153-4162.
23. Xu, M., Zhang, L., Du, B., et al., "The linear mixed model constrained particle swarm optimization for hyperspectral endmember extraction from highly mixed data". In: *2016 8th Workshop on Hyperspectral Image and Signal Processing -- Evolution in Remote Sensing*, Los Angeles, USA: IEEE, 2016:1-4.
24. Kubokawa, T., "Constrained empirical Bayes estimator and its uncertainty in normal linear mixed models". *Journal of Multivariate Analysis*, 122, 2013, pp. 377-392.
25. Hoyer, P. O., "Non-negative matrix factorization with sparseness constraints". *Journal of Machine Learning Research*, 5(11), 2004, pp. 1457-1469.
26. Jia, S., Qian, Y., "Constrained non-negative matrix factorization for hyperspectral unmixing". *IEEE Transactions on Geoscience and Remote Sensing*, 47(1), 2009, pp. 161-173.
27. Yokoya, N., Yairi, T., Iwasaki, A., "Coupled non-negative matrix factorization unmixing for hyperspectral and multispectral data fusion". *IEEE Transactions on Geoscience and Remote Sensing*, 50(2), 2012, pp. 528-537.
28. Chen, Q., Shi, Z., An, Z., "Hyperspectral image fusion based on sparse constraint NMF". *International Journal for Light and Electron Optics*, 125(2), 2014, pp. 832-838.
29. Paatero, P., Tapper, U., "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values". *Environmetrics*, 5(2), 1994, pp. 111-126.
30. Cheng, M., S., "A practical fast NMF algorithm". *Journal of Dalian University of Technology*, 53(1), 2013, pp. 151-156.
31. U.S. Army Geospatial Center, <http://www.agc.army.mil/>.
32. Leica Geosystems Geographic Imaging web site, <http://gi.leica-geosystems.com>.