Hyperspectral image fusion based on sparse constraint NMF

Quan Chen, Zhenwei Shi [*], Zhenyu An

Image Processing Center, School of Astronautics, Beihang University, Beijing 100191, P.R. China

Abstract

The spatial resolution of hyperspectral image is often low due to the limitation of the imaging spectrometer. Fusing the original hyperspectral image with high-spatial-resolution panchromatic image is an effective approach to enhance the resolution of hyperspectral image. However, it is hard to preserve the spectral information at the same time of enhancing the resolution by the traditional fusion methods. In this paper, we proposed a fusion method based on the spectral unmixing model called sparse constraint nonnegative matrix factorization (SCNMF). This method has a superior balance of the spectral preservation and the spatial enhancement over some traditional fusion methods. In addition, the added sparse prior and NMF based unmixing model make the fusion more stable and physically reasonable. This method first decomposes the hyperspectral image into an endmember-matrix and an abundance-matrix, then sharpens the abundance-matrix with the panchromatic image, finally obtains the fused image by solving the spectral constraint optimization problem. The experiments on both synthetic and real data show the effectiveness of the proposed method.

Key words: Hyperspectral image fusion; Resolution enhancement; Nonnegative matrix factorization(NMF); Sparse prior.

1 Introduction

Hyperspectral remote sensing technology has been successfully applied in areas such as anomaly detection, target recognition and background characterization due to its high resolution spectral dimension [25]. However, for the limitation of the imaging spectrometer, the spatial resolution of hyperspectral image (HSI) is often low, which hinders the further improvement of many applications [4]. Compared with HSI, panchromatic image (PI) has much higher spatial resolution in the same scene. So many scholars have been proposing various methods to fuse the HSI and PI to obtain one high-spatial-and-spectral resolution image.

In general, the most common fusion algorithms can be separated into three categories [4]. The most classical one is the projection and substitution based methods, which assume that the PI is equivalent to the structural component of the HSI when translated the HSI into a new space [16]. Algorithms like Intensity-Hue-Saturation (IHS) [3, 7, 27], Principal Component Substitution (PCS) [6] and Gram Schmidt Transformation (GST) [15] are all this kind of methods. Another kind of methods are based on band ratio and arithmetic combinations, such as Synthetic Variable Ratio (SVR) [20], which performs well and time-consuming is small. The last category is the wavelet based methods, such as Discrete Wavelet Transform (DWT) [13, 22], which use the DWT to extract the high frequencies of the PI and then inject them into the HSI to get the fused image. Of course, in resent years, there are many other algorithms like Fast Fourier Transform enhanced IHS [14] are proposed to improve the existing methods. What's more, Li et al. [16] introduce the compressed sensing technique into the fusion of multispectral image (MSI) and PI, which gives a new approach to image fusion.

As we will see latter, all of the methods mentioned before have their shortcomings for the fusion of HSI and PI. HSI, compared with MSI and PI, has much more abundant spectral information, and many applications such as classification are base on this fact. Therefore, it is worth nothing that on the processing of image fusion, we must preserve the spectral information and enhance the spatial resolution at the same time. Unfortunately, most of the fusion methods like PCS and DWT, which perform well for the fusion of MSI and PI, can not achieve this directly. It is clear that one hyperspectral image only contains several materials, which we call each of them endmember in the field of hyperspectral image processing. Form this point of view, the fused high-spatial-resolution image must have the same endmembers with the original low-spatial-resolution image. However, none of the methods mentioned before has the theoretically guarantee. Some researchers address this problem from the perspective of spectral unmixing [9, 24, 28–31], which we named them as spectral unmixing based fusion methods. These methods share the

same fundamental steps, that is, first unmix the low-spatial-resolution HSI into a endmember-matrix and abundance-matrix, then fuse the abundance-matrix with the high-spatial-resolution PI using constrained optimization techniques. With the constant development of hyperspectral unmixing technique, the idea of using unmixing model for the fusion of HSI and PI is not only easy to implement but also physically reasonable.

In the past few decades, researchers have proposed a large number of algorithms for hyperspectral unmixing based on linear mixing model (LMM) [19]. Among these methods, nonnegative matrix factorization (NMF) [18] has become a fairly useful hyperspectral unmixing method, for its mathematical form and constraints are in accordance with the physically meaning of hyperspectral unmixing. NMF, which is reported as simulating how the brain identify objects, originally attempts to learn a parts-based representation of data. In brief, given a nonnegative data matrix $\mathbf{V}$, NMF algorithm seeks two nonnegative matrix $\mathbf{W}$ and $\mathbf{H}$, which satisfy the equation $\mathbf{V} \approx \mathbf{WH}$. When NMF is introduced to hyperspectral unmixing, the matrix $\mathbf{V}$, $\mathbf{W}$ and $\mathbf{H}$ represent the original low-spatial-resolution HSI, endmember-matrix and abundance-matrix, respectively. As the nonconvexity of the objective function derived from the NMF decomposition model, there are a large number of local minima. So it is impossible to obtain a global optimal solution. However, we can introduce some priori information to make the solution more reasonable. Hoyer [10,11] introduces the sparse priori into NMF and gets a combination of sparse coding and NMF. Noticing that for every pixel of HSI, it contains only a few of endmembers extracted form the whole image, that is, the abundance fraction for each pixel is sparse. inspired by this, Qian et al. [23] extend the NMF method by incorporating the $L_{1/2}$ sparsity constraint, which named as $L_{1/2} - \mathbf{NMF}$, and successfully apply it to hyperspectral unmixing.

Among these spectral unmixing based fusion methods, [9, 24] require a proper spectral library or some prior about the content of the image, which limit their application. What's more, on the sharpening process, they also require solve a underdetermined problem, which leads the abundance fractions' mapping error, and further the spectral distortion of the fused HSI. [28,29] use Coupled Nonnegative Matrix Factorization (CNMF) to unmixing the original low-spatial-resolution HSI and the original high-spatial-resolution MSI to obtain two couple of endmember-matrixes and abundance-matrixes. Then by combining the endmember-matrix obtained from HSI and the high-spatial-resolution abundance-matrix obtained from MSI, the fused high-spatial-resolution HSI can be generated. However, as the point spread function (PSF) and spectral response function (SRF), which are used to model the spatial resolution degradation from the original source MSI to the original source HSI and the spectral resolution degradation from latent HSI (in fact, it can also be substituted by original source HSI) to original source MSI, respectively, are hard to estimated in the real data, CNMF method is hard to apply to real HSI. Furthermore,

3

if PI is used as the high-spatial-resolution source for HSI fusion, it is hard to use NMF to decompose a endmember-matrix and abundance-matrix from it. [30, 31] have introduced a spectral preservation constraint to the NMF-based fusion model, however, it does not explore the inherent sparse priori of the hyperspectral data's abundance-matrix during the decomposition process.

Inspired by the recently hyperspectral unmixing methods and the spectral unmixing based fusion model, we proposed a sparse constrained nonnegative matrix factorization (SCNMF) fusion method. This method first uses the recently well performed unmixing algorithm to give a better hyperspectral endmember-matrix and abundance-matrix, then sharpens the abundance-matrix with the PI, finally produces the fused HSI with the spectral preservation constraint. The proposed method has three advantages compared with other HSI fusion methods: (1) it preserves the original HSI endmembers and gives a more reasonable physical explanation; (2) it presents a direct spectral preservation constraint that make sure the fused HSI has little spectral distortion; (3) it does not need PSF and SRF estimation, which makes the application to real HSI become possible.

The remainder of this paper is organized as follows. Section 2 gives a brief introduction of sparse NMF and explain its meaning in the fusion model. Section 3 describes the proposed method for the fusion of HSI and PI, and derives the updata algorithm. Section 4 presents the experimental results on synthetic data and real data, and gives the corresponding discussion. At last, section 5 concludes this paper.

2 Sparse NMF

Since Lee et al. [18] proposed NMF, it has attracted many researchers' interests and has been applied to kinds of fields because of its simplicity of mathematical form and easy to implement. For a standard NMF problem, given a nonnegative data matrix $\mathbf{V} \in \mathbb{R}^{L \times K}$, NMF algorithm seeks two nonnegative matrix $\mathbf{W} \in \mathbb{R}^{L \times S}$ and $\mathbf{H} \in \mathbb{R}^{S \times K}$ which satisfy:

$$\mathbf{V} \approx \mathbf{WH} \tag{1}$$

To obtain the decomposed matrix $\mathbf{W}$ and $\mathbf{H}$, standard NMF algorithm simply minimizes the difference between $\mathbf{V}$ and $\mathbf{WH}$ using Euclidean distance. We display the cost function as follows:

$$C(\mathbf{W}, \mathbf{H}) = \frac{1}{2} \sum_{j=1}^{K} \|\mathbf{V}_j - \mathbf{W}\mathbf{H}_j\|_2^2 = \frac{1}{2} \|\mathbf{V} - \mathbf{W}\mathbf{H}\|_F^2 \qquad (2)$$

However, as discussed in section 1, because of the nonconvexity of the cost function $C(\mathbf{W}, \mathbf{H})$ with respect to both $\mathbf{W}$ and $\mathbf{H}$, it is impossible to obtain the global optimal solution. In order to make the solution more stable, we can add some reasonable prior to the cost function depending on the problem. As for HSI decomposition, sparsity is a reasonable constraint.

When NMF based method has been applied to hyperspectral image fusion, the mathematical symbol in equation (2) is explained as follows: $\mathbf{V}$ denotes the original low-spatial-resolution HSI which contains L bands and K pixels, that is, each column of $\mathbf{V}$, denoted as $\{\mathbf{V}_j\}_{j=1}^{K} \in \mathbb{R}^L$, stands for one pixel with L bands. $\mathbf{W}$ is the endmember-matrix, each column of which represents one endmember spectrum (the spectrum of a pure material, such as water, tree, soil, etc). $\mathbf{H}$ is the so called abundance-matrix, each column of which denoted as $\{\mathbf{H}_j\}_{j=1}^{K} \in \mathbb{R}^S$, and represents the fractions of all the endmembers in the $\mathbf{W}$ for one pixel in $\mathbf{V}$. As the endmembers $\mathbf{W}$ is extracted from the whole HSI, for any pixel $\mathbf{V}_j$, its fractions $\mathbf{H}_j$ should be sparse. In other words, we can reconstruct each pixel in HSI efficiently with only a few of endmembers in $\mathbf{W}$. For this reason, we add the sparse constraint to (2) to obtain the following sparse NNF objective fuction:

$$C(\mathbf{W}, \mathbf{H}) = \frac{1}{2} \|\mathbf{V} - \mathbf{W}\mathbf{H}\|_F^2 + \alpha \sum_{j=1}^{K} \|\mathbf{H}_j\|_1 \qquad (3)$$

where

$$\|\mathbf{H}_j\|_1 = \sum_{i=1}^{S} |\mathbf{H}_{ij}| \qquad (4)$$

is the vector 1-norm which is used to measure the sparsity of vector $\mathbf{H}_j$. Based on the multiplicative rule described in [18], Hoyer [10] proposed a update method for solving the sparse NMF described in (3) as follows:

$$\mathbf{W} \leftarrow \mathbf{W} .* \mathbf{V}\mathbf{H}^T ./ \mathbf{W}\mathbf{H}\mathbf{H}^T \qquad (5)$$

$$\mathbf{H} \leftarrow \mathbf{H} .* \mathbf{W}^T\mathbf{V} ./ (\mathbf{W}^T\mathbf{W}\mathbf{H} + \alpha\mathbf{1}) \qquad (6)$$

where $\mathbf{1}$ denotes a matrix whose elements are all 1, (.*) and (./) denote the elementwise multiplication and division, respectively. $(\triangle)^T$ denotes the transpose of the matrix $\triangle$. Of course, some researches [23] use the so called

$L_{1/2}$ sparse measure in lieu of the L_1, however, in fact, they share the same fundamental idea.

3 Fusion model base on SCNMF

3.1 The proposed model

As mentioned earlier, in order to obtain a fused high-spatial-resolution HSI with a little spectral distortion, the fusion model must satisfy the following two conditions: (1) the sharpening information (or high-frequency information) extracted from PI must be injected into the original low-spatial-resolution HSI properly; (2) there should be some apparent spectral preservation measure to guarantee the fused image's spectrum approximates the original HSI's as much as possible for every pixel. From the aspect of fusing image based on the unmixing model, we discuss the proposed SCNMF fusion method in detail as follows:

According to the condition (1), the original low-spatial-resolution HSI $\mathbf{V}$ can be decomposed into the endmember-matrix $\mathbf{W}$ and abundance-matrix $\mathbf{H}$, which has been discussed in detail in section 2. It is obvious that $\mathbf{H}$ contains the low-resolution information of $\mathbf{V}$. Therefore, it is reasonable to obtain a high-spatial-resolution HSI by the way of sharping $\mathbf{H}$ with the high-resolution details in PI. The proposed fusion process is illustrated in Fig. 1. Thus, we can express it as this optimization problem:

$$
\begin{aligned}
min \quad & C(\mathbf{W}, \mathbf{H}) = \frac{1}{2}\|\mathbf{V} - \mathbf{WH}\|_F^2 + \alpha \sum_{j=1}^{K} \|\mathbf{H}_j\|_1 \\
s.t. \quad & \mathbf{W} \geq 0, \mathbf{H} \geq 0
\end{aligned}
\tag{7}
$$

And then we have the fused HSI $\mathbf{V}_f$:

$$
\mathbf{V}_f = \mathbf{W}(\beta \mathbf{H} + (1 - \beta)\mathbf{P})
\tag{8}
$$

where $\mathbf{P} \in \mathbb{R}^{S \times K}$, each of whose row denoted as $\{\mathbf{P}_i\}_{i=1}^{S} \in \mathbb{R}^K$ is the vectorized PI, represents the replication of the PI. Parameter β balances the abundance preservation of the original HSI and the degree of sharpening.

However, if we simply have the fused HSI from formula (8) after solving the optimization problem (7), it is obvious that this manipulation will lead

to serious spectral distortion. According to the condition (2), we derive a apparent spectral preservation term in this paper.

On the process of HSI fusion, we always hope that the fused HSI have the same spectral characteristics with the original HSI, that is, for every pixel, the fused one and the original one have the same trend of spectral curve. Suppose $\{\mathbf{V}_j\}_{j=1}^{K} \in \mathbb{R}^L$ and $\{\mathbf{V}_{fj}\}_{j=1}^{K} \in \mathbb{R}^L$ denote the original HSI and fused HSI pixel, respectively. We can simply minimize the following formula to preserve the spectral characteristics:

$$\mathcal{S}(\mathbf{V}, \mathbf{V}_f) = \sum_{i=1}^{K}(\|\mathbf{V}_i\|^2\|\mathbf{V}_{fi}\|^2 - (<\mathbf{V}_i, \mathbf{V}_{fi}>)^2) \qquad (9)$$

where $\|\triangle\|$ and $<\triangle>$ denote the Euclidian L_2 norm and the inner product. Borrowing the representation of matrix and some matrix simplification tricks, we can simply (9) to (10)

$$\mathcal{S}(\mathbf{V}, \mathbf{V}_f) = tr((\mathbf{V}_f^T\mathbf{V}_f).*(\mathbf{V}^T\mathbf{V})) - tr((\mathbf{V}_f^T\mathbf{V}).*(\mathbf{V}_f^T\mathbf{V})) \qquad (10)$$

where $tr(\triangle)$ denotes the trace of the matrix $\triangle$. From formula (9) or (10), it is clear that the better the spectral preserved, the smaller the value will be. In the ideal case, that is, there is no spectral distortion form $\mathbf{V}$ to $\mathbf{V}_f$, the value in formula (10) will be 0.

When we add the spectral preservation formula (10) to the spectral un-mixing based model expressed in formula (7), we obtain the final fusion model as follows:

$$min \quad F(\mathbf{W}, \mathbf{H}) = \frac{1}{2}\|\mathbf{V} - \mathbf{W}\mathbf{H}\|_F^2 + \alpha\sum_{j=1}^{K}\|\mathbf{H}_j\|_1 + \gamma\mathcal{S}(\mathbf{V}, \mathbf{V}_f)$$
$$s.t. \quad \mathbf{W} \geq 0, \mathbf{H} \geq 0 \qquad (11)$$

where γ is a parameter that balance the spectral quality and the spatial quality of the fused HSI. Finally, we have our fused HSI $\mathbf{V}_f$ expressed in the earlier mentioned formula (8).

3.2 *Updating rules*

Compared with the solution of the standard NMF or sparse NMF model described in [18] and [10], respectively, the result of the proposed SCNMF
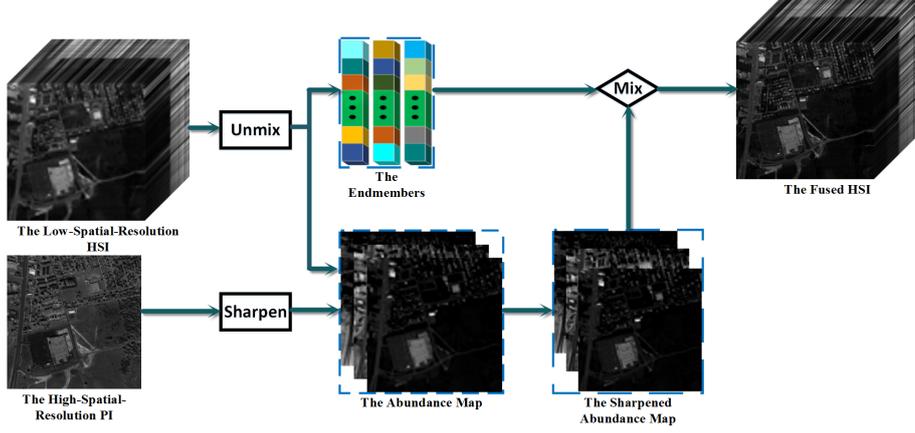
Fig. 1. The proposed fusion process

fusion model in formula (11) is much more stable. However, it can't change the fact that the cost function $F(\mathbf{W}, \mathbf{H})$ in formula (11) is still not convex with respect to both $\mathbf{W}$ and $\mathbf{H}$, which leads that finding the global minimum is still not realistic. Here, we derive a new multiplicative update rule to minimize the objective cost function $F(\mathbf{W}, \mathbf{H})$.

In order to obtain the multiplication factors with respective to $\mathbf{W}$ and $\mathbf{H}$, we first take the partial derivative with respect to $\mathbf{W}$ and $\mathbf{H}$ and get:

$$
\begin{aligned}
\frac{\partial F(\mathbf{W}, \mathbf{H})}{\partial \mathbf{W}} =& (\mathbf{W}\mathbf{H} - \mathbf{V})\mathbf{H}^T + 2\gamma \mathbf{W}\mathbf{H}_0 diag(\mathbf{V}^T\mathbf{V})\mathbf{H}_0^T - \\
& 2\gamma \mathbf{V} diag(\mathbf{H}_0^T\mathbf{W}^T\mathbf{V})\mathbf{H}_0^T
\end{aligned}
\tag{12}
$$

$$
\begin{aligned}
\frac{\partial F(\mathbf{W}, \mathbf{H})}{\partial \mathbf{H}} =& \mathbf{W}^T(\mathbf{W}\mathbf{H} - \mathbf{V}) + 2\gamma\beta^2\mathbf{W}^T\mathbf{W}\mathbf{H} diag(\mathbf{V}^T\mathbf{V}) + \\
& 2\gamma\beta(1 - \beta)\mathbf{W}^T\mathbf{W}\mathbf{P} diag(\mathbf{V}^T\mathbf{V}) - \\
& 2\gamma\beta^2\mathbf{W}^T\mathbf{V} diag(\mathbf{H}^T\mathbf{W}^T\mathbf{V}) - \\
& 2\gamma(1 - \beta)^2\mathbf{W}^T\mathbf{V} diag(\mathbf{P}^T\mathbf{W}^T\mathbf{V}) + \alpha\mathbf{1}
\end{aligned}
\tag{13}
$$

where $\mathbf{H}_0 = \beta\mathbf{H} + (1 - \beta)\mathbf{P}$ and $diag(\triangle)$ is a diagonal matrix whose diagonal elements are the corresponding diagonal ones from matrix $\triangle$. Then we can easily get the multiplication factor with respect to $\mathbf{W}$ by taking positive and negative terms in formula (12) (the partial derivative of $F(\mathbf{W}, \mathbf{H})$ with respect to $\mathbf{W}$) as its denominator and numerator respectively, and so does the multiplication factor with respect to $\mathbf{H}$. Following this principle, the updating rules can be acquired as:

$$\mathbf{W} \leftarrow \mathbf{W}.*(\mathbf{V}\mathbf{H}^T + 2\gamma\mathbf{V}diag(\mathbf{H}_0^T\mathbf{W}^T\mathbf{V})\mathbf{H}_0^T)./$$
$$(\mathbf{W}\mathbf{H}\mathbf{H}^T + 2\gamma\mathbf{W}\mathbf{H}_0 diag(\mathbf{V}^T\mathbf{V})\mathbf{H}_0^T) \tag{14}$$

$$\mathbf{H} \leftarrow \mathbf{H}.*(\mathbf{W}^T\mathbf{V} + 2\gamma\beta^2\mathbf{W}^T\mathbf{W}\mathbf{P}diag(\mathbf{V}^T\mathbf{V})+$$
$$2\gamma\beta^2\mathbf{W}^T\mathbf{V}diag(\mathbf{H}^T\mathbf{W}^T\mathbf{V})+$$
$$2\gamma(1-\beta^2)\mathbf{W}^T\mathbf{V}diag(\mathbf{P}^T\mathbf{W}^T\mathbf{V}))./$$
$$(\mathbf{W}^T\mathbf{W}\mathbf{H} + 2\gamma\beta^2\mathbf{W}^T\mathbf{W}\mathbf{H}diag(\mathbf{V}^T\mathbf{V})+ \tag{15}$$
$$2\gamma\beta\mathbf{W}^T\mathbf{W}\mathbf{P}diag(\mathbf{V}^T\mathbf{V}) + \alpha\mathbf{1})$$

3.3 Implementation issues

There are several important implementation issues that should call our attention. The first one is about the initialization methods of the endmember-matrix $\mathbf{W}$ and the abundance-matrix $\mathbf{H}$. Since Solving the optimization problem described by (11) often results in a local minimum, different initialization methods not only affect the convergence rate of the algorithm, but result different solutions. We can simply choose values between 0 and 1 randomly as the entries of $\mathbf{W}$ and $\mathbf{H}$, however, in this paper, vertex component analysis (VCA) [21] are selected to make the initialization of the SCNMF fusion algorithm.

The second important issue is how to use the PI to sharpen the abundance-matrix $\mathbf{H}$. We can directly use the PI to generate the sharpening matrix $\mathbf{P}$ and substitute into formula (8) to obtain the fused image. However, if we do so, the sharpening matrix $\mathbf{P}$ wil contain much low frequency information of PI, which will lead to serious spectral distortion at the same time of sharpen the abundance-matrix $\mathbf{H}$. Thus, in the actual implementation, we employ high frequency filter to process PI before we use it to generate matrix $\mathbf{P}$.

The third one is about the parameters we choose for the algorithm. There are three parameters in SCNMF: α, β and γ are the controller for the degree of sparse, sharpening and spectral preservation, respectively. Different parameters selection scheme will have a different result. In the simulated and real experiments carried out in the next section, we both set them as $\alpha = 0.01, \beta = 0.4, \gamma = 0.01$.

The stopping criteria what we choose is also a key issue. Here, a predefined value, denoted as **tol**, is set as the relative error tolerance. If the relative error (**Rerror**) of the cost function $F(\mathbf{W}, \mathbf{H})$ defined as follows less than **tol** or

the number of iteration exceeds the maximum iteration number, the iteration ends.

$$\mathbf{Rerror} = |\frac{F(\mathbf{W}, \mathbf{H})_{new} - F(\mathbf{W}, \mathbf{H})_{old}}{F(\mathbf{W}, \mathbf{H})_{old}}| \leq \boldsymbol{tol} \qquad (16)$$

where $F(\mathbf{W}, \mathbf{H})_{old}$ and $F(\mathbf{W}, \mathbf{H})_{new}$ denotes the cost function's value of the last and current iteration, respectively.

The proposed SCNMF fusion approach for HSI is summarized in Algorithm 1.

SCNMF for fusion of HSI and PI

---

1. **Input:** HSI data $\mathbf{V} \in \mathbb{R}^{L \times K}$ and sharpen matrix $\mathbf{P} \in \mathbb{R}^{S \times K}$ generated form PI.

2. **Initialize W** and **H** by VCA.

3. **Repeat** until convergence:
   a) Update **W** by (14)
   b) Update **H** by (15)

4. **Output:** obtain the fused HSI by (8).

---

4   Experiments

In this section, we first give some common evaluation criteria to measure the results of fusion objectively. Then we move on to carry out experiments on both synthetic HSI and real HSI, which we get from the internet [12], to demonstrate the proposed SCNMF fusion method.

4.1 Evaluation criteria

To demonstrate the effectiveness of the proposed SCNMF fusion method, we need to compare the fusion result with other fusion methods quantitatively. However, quality assessment of the fused image is a much-debated problem since a unique, reliable image quality index is unavailable today yet [4]. Therefore, it requires computation of a number of different indices with the fused HSI, the original HSI and the reference HSI [2]. In this paper, the following typical evaluation metrics are used.

(1) the Spectral Angle Mapper (**SAM**) denotes the absolute value of the spectral angle which reflects the degree of the spectral distortion from the original HSI $\mathbf{V}$ to the fused HSI $\mathbf{V}_f$. Let column vector $\mathbf{V}_i$ and $\mathbf{V}_{fi}$ denote the spectral vector of a pixel of $\mathbf{V}$ and $\mathbf{V}_f$ respectively. Then the **SAM** of this pixel can be define as:

$$\mathbf{SAM} = \arccos(\frac{<\mathbf{V}_i, \mathbf{Vf}_i>}{\|\mathbf{V}_i\|_2 \|\mathbf{V}_{fi}\|_2}) \tag{17}$$

The final **SAM** is averaged over the whole image to yield a global measurement of spectral distortion [2]. In the ideal case, the **SAM** would be 0, that is, there is no spectral distortion.

(2) the Average Gradient (**AG**) can be used to measure the sharpness of the image, that is, the richer of the details in the image, the greater **AG** will be. It is defined as:

$$\mathbf{AG} = \frac{1}{(M-1)(N-1)} \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{\frac{(\nabla x)^2 + (\nabla y)^2}{2}} \tag{18}$$

where M and N are the height and width of the image, ∇x and ∇y are the gradient of the x and y directions.

(3) the Correction Coefficient (**CC**) can be calculated by:

$$\mathbf{CC} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} [H(i,j) - \overline{H}][H_f(i,j) - \overline{H_f}]}{\sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} [H(i,j) - \overline{H}] \sum_{i=1}^{M} \sum_{j=1}^{N} [H_f(i,j) - \overline{H_f}]}} \tag{19}$$

where H and H_f denote one band of the original HSI and the fused HSI with size $M \times N$, respectively. $\overline{H}$ and $\overline{H_f}$ denote the average of one band of the original HSI and the fused HSI, respectively. We get the final **CC** by average over the whole HSI spectral axis. **CC** reflects the correlation between the original HSI and the fused HSI.

(4) the Spectral Information Divergence (SID) is a spectral similarity measure to capture the spectral correlation between the original HSI pixel and the fused HSI pixel [5]. Assume that $\mathbf{x}=(x_1,\ldots,x_L)$ and $\mathbf{x}=(x_1,\ldots,x_L)$ are pixels taken from the original HSI and the fused HSI. Then it can be defined as:

$$\mathbf{SID}(\mathbf{x},\mathbf{y}) = D(\mathbf{x}\|\mathbf{y}) + D(\mathbf{y}\|\mathbf{x}) \tag{20}$$

where $D(\mathbf{x}\|\mathbf{y})$ is the so called cross-entropy that means the relative entropy of $\mathbf{y}$ with respect to $\mathbf{x}$ and is given by:

$$D(\mathbf{y}\|\mathbf{x}) = \sum_{l=1}^{L} p_l \log_2(\frac{p_l}{q_l}) \tag{21}$$

where p_l and q_l are calculated by:

$$p_l = \frac{x_l}{\sum_{i=1}^{L} x_i} \tag{22}$$

$$q_l = \frac{y_l}{\sum_{i=1}^{L} y_i} \tag{23}$$

The final **SID** is also given by averaging over the total pixels from the HSI.

(5) the Root Mean Square Error (**RMSE**) is a common index to measure the error between two images. one band of HSI's $\mathbf{RMSE}_k$ is defined as:

$$\mathbf{RMSE}_k = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [H_f(i,j) - H_r(i,j)]^2} \tag{24}$$

where H_f and H_r represent one band of the fused HSI and the reference HSI, respectively. We get the final **RMSE** by averaging over the whole band of the HSI.

(6) the relative global synthesis error (**ERGAS**) is defined as:

$$\mathbf{ERGAS} = 100\frac{h}{l}\sqrt{\frac{1}{L}\sum_{k=1}^{L}\frac{\mathbf{RMSE}_k^2}{\mathbf{MEAN}_k^2}} \tag{25}$$

where h and l denote as the spatial resolution of PI and HSI, respectively. L is the number of band of the original HSI. $\mathbf{MEAN}_k$ is the mean value of the kth band of the reference HSI. $\mathbf{RMSE}_k$ is the **RMSE** of the kth band between the fused HSI and the reference HSI , which is defined by formula (24). It is obvious that smaller **ERGAS** suggest the better fusion result.
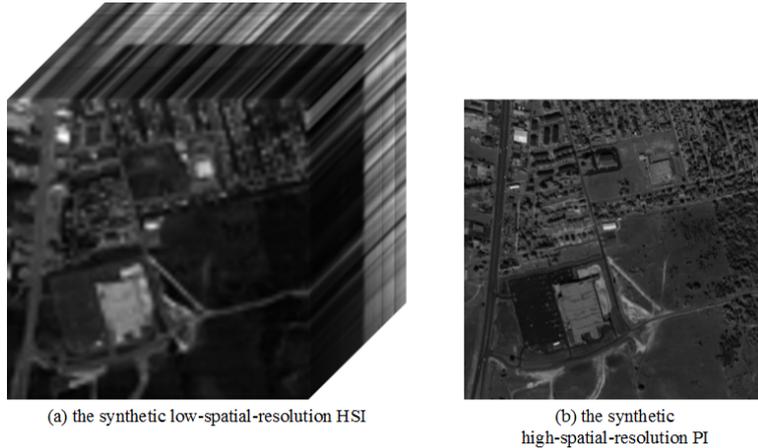
(a) the synthetic low-spatial-resolution HSI   (b) the synthetic high-spatial-resolution PI

Fig. 2. the schematic diagram of the Synthetic HSI and PI

4.2   Experiment on Synthetic Data

As we know, we have to carry out registration between HSI and PI before we fused them. However, misregistration of the source images (HSI and PI) can cause the generation of mismatched fused image [17], which will have some effect on the evaluation of the performance of different methods. In addition, some indices such as **ERGAS** need the reference HSI to give the objective evaluation. For these reason, we download the HSI of the Northwest Tippecanoe County, Indiana from [12] and use it to carry out a simulated experiment.

The HSI from [12] contains 220 bands with the size of 307×306, which comes from the AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) built by JPL. We first remove the low-SNR bands and then randomly select 100 bands to generate the reference HSI and one band to simulate the high-spatial-resolution PI. Next, we simulate the original low-spatial-resolution HSI by downsampling the reference HSI with the factor of 4 and resizing (bicubic interpolation) it to the same size as before. Thus, we get the source images which have been geometricly registrated with high accuracy, which are shown in Fig 2.

The original HSI, the reference HSI and the fused results obtained form different methods shown in Fig 3 are composed of three bands from the corresponding HSI. From the aspect of visual analysis, we mainly focus on two important issues, that is, the spectral preservation and the spatial enhancement. It is clear that in Fig 3, Compared with the fusion result obtained from the proposed fusion method, the fusion results from DWT and GST have much more serious color distortion (In the HSI, we often call them spectral distortion) based on the reference HSI. The result form Ehlers [8] has a obvious blur effects. It is seem that in visual effects, high-pass filtering (HPF) [1] and
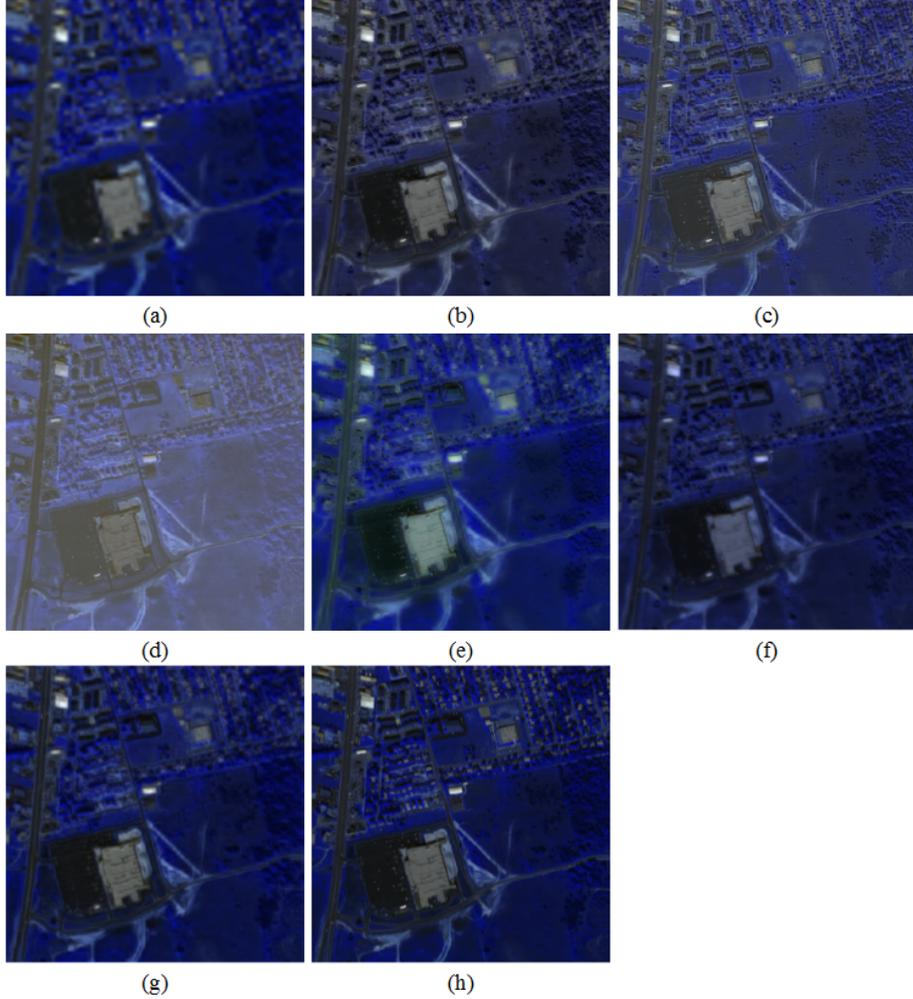
Fig. 3. The fused results from different methods in simulated experiment. (a)The resampled low-spatial-resolution HSI; (b)-(g) are the fused results from DWT, HPF, GST, Ehlers, SPNMF and the proposed method; (h)the reference HSI.

SPNMF [30] have the same fusion result with the proposed method (in fact, if you observe carefully, HPF has more color distortion and SPNMF has more blur effects). In order to give a more convincing comparative result, we show the objective indices for these methods in table 1.

From table 1, we can see that whether form the aspect of spectral preservation or the spatial enhancement, the proposed fusion method has a better performance than DWT, Ehlers and SPNMF. However, the proposed method's **AG** is smaller than those of HPF and GST (but very close). This result suggest that the degree of spatial enhancement of the proposed method is not as good as that of HPF and GST. But as for other indices, such as SAM, the proposed method's are much better than GST's. What's more, the fusion result of GST in Fig. 3 show that it has a serious spectral distortion. Now let us analyse the comparison of HPF and the proposed method. Though HPF

Table 1
Evaluation results of the simulated experiment

	DWT	HPF	Ehlers	GST	SPNMF	The proposed SCNMF method
SAM	11.7628	5.5081	6.7069	26.5828	6.0229	**1.4754**
AG	0.0214	**0.0276**	0.0214	0.0275	0.0197	0.0271
CC	0.8976	0.8854	0.9359	0.6652	0.8978	**0.9383**
RMSE	0.9634	1.2707	6.9038	1.4881	4.1634	**0.0614**
ERGAS	11.6152	7.915	42.0151	8.4659	26.9675	**0.4036**
SID	0.0841	0.0624	0.0146	0.2217	0.0252	**0.0013**

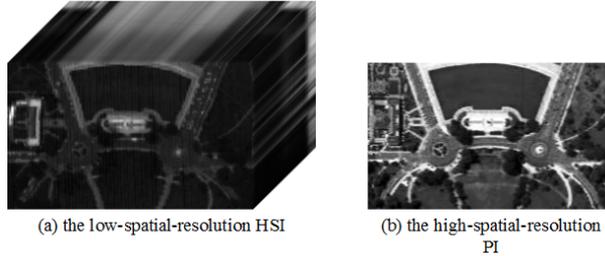(a) the low-spatial-resolution HSI (b) the high-spatial-resolution PI

Fig. 4. the schematic diagram of the real HSI and PI.

show a better sharpening than the proposed method, the larger indices such as **SAM** and **SID** suggest that the distortion of HPF is more serious than that of the proposed method, which has a important effect in many hyperspectral applications. In addition, the smaller **RMSE** and **ERGAS** suggests that the fusion result of the proposed method approximates the reference HSI much more better. The phenomenon of larger **AG** but larger **RMSE** and **ERGAS** implies that **HPF** may introduce some artifacts in the fused HSI. So in general, the propose method do a better balance between spatial enhancement and spectral preservation.

4.3 Experiment on real Data

We also carry out the fusion experiment on the real data. We first download the HSI and PI (the spatial resolution is twice higher than HSI) of Washington DC Mall from [12] and google map, respectively. After removing low-SNR bands, We then carry out registration with them and finally get the PI and the original low-spatial-resolution HSI with 64 bands and the size of 150×250. The image used to fuse are shown in Fig. 4.

The original low-spatial-resolution HSI and the fused results obtained form different methods shown in Fig. 5 are composed of three bands from the corresponding HSI. It shows that in the original low-spatial-resolution HSI,
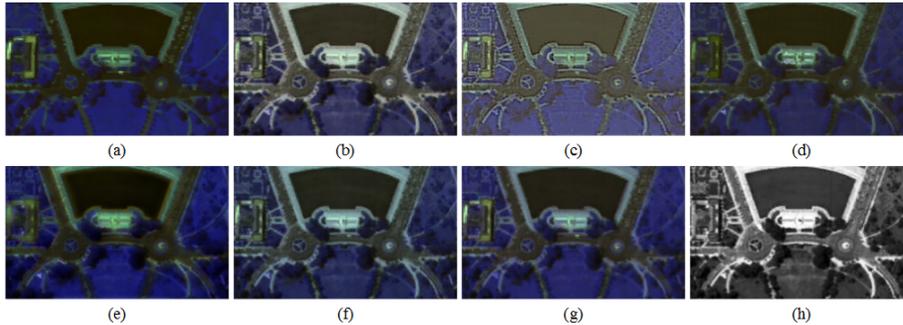
Fig. 5. The fused results from different methods in real experiment. (a)The low-s-patial-resolution HSI; (b)-(g) are the fused results from DWT, HPF, GST, Ehlers, SPNMF and the proposed method; (h)the high-spatial-resolution PI.

Table 2
Evaluation results of the real experiment

	DWT	HPF	Ehlers	GST	SPNMF	The proposed SCNMF method
SAM	14.5275	20.2814	13.8341	30.6976	12.9038	**8.9042**
AG	**0.0363**	0.0358	0.0221	0.0203	0.0308	0.0297
CC	0.7333	0.7117	0.7977	0.7947	0.7583	**0.8868**
SID	0.082	0.1036	0.0659	0.0882	0.074	**0.0423**

many details, such as some roofs and roads, have been disappeared. However, after fusion, these details have been added to the fused image in some degree. The DWT and HPF enhance the spatial resolution better than other methods, however, both of DWT and HPF have serious spectral distortion. Relatively speaking, the SPNMF and the proposed methods have the best result. Compared with the SPNMF, since the sparse prior have been added, the proposed method makes the spectral unmixing more stable and reasonable, further makes the spectral distortion much less. In this HSI fusion, the **AG** of the propose method is a little less than that of SPNMF, however, the proposed method makes a much progress in the aspect of spectral preservation. These results suggest that the proposed method can also do a better balance between the spectral preservation and the spatial enhancement in the real data.

5   Conclusions

In this paper, a new hyperspectral image fusion algorithm called sparse constraint nonnegative matrix factorization (SCNMF) has been proposed. This method is based on spectral unmixing model. The algorithm first utilizes the sparse NMF to decompose the original hyperspectral image to an endermember-matrix and an abundance-matrix, then sharpens the abundance-

matrix by the high-spatial-resolution panchromatic image, and finally produce the sharpened hyperspectral image with the spectral preservation constraint. Unlike some spectral unmixing based methods, such as CNMF, can not apply to the fusion of real hyperspectral image and panchromatic image, the proposed method need not any prior knowledge about the content of the image or PSF and SRF. Thus it can be applied to the synthetic or real hyperspectral fusion of both multispectral image and panchromatic image. Our experiment show that the proposed method can do a better balance between the spectral preservation and resolution enhancement than some traditional methods.

References

[1] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, Context driven fusion of high spatial and spectral resolution images based on oversampled multi-resolution analysis, IEEE Transaction on Geoscience and Remote Sensing 40(10) (2002) 2300C2312.

[2] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, Comparison of pansharpening algorithms: outcome of the 2006 GRS-S data-fusion contest, IEEE Transaction on Geoscience and Remote Sensing 45(10) (2007) 3012-3021.

[3] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, The use of intensity hue saturation transformations for merging SPOT panchromatic and multispectral image data, Photogrammetric Engineering and Remote Sensing 56(4) (1990) 459-467.

[4] M. Cetin, and N. Musaoglu, Merging hyperspectral and panchromatic image data: qualitative and quantitative analysis, International Journal of Remote Sensing 30(7) (2009) 1779-1804.

[5] C-I. Chang, An information-theoretic approach to spectralVariability, similarity, and discrimination for hyperspectral image. IEEE Transaction on Information Theory 46(5) (2000) 1927-1932.

[6] P. S. Chavez, S. C. Sides, and J. A. Anderson, Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic, Photogrammetric Engineering and Remote Sensing 57(3) (1991) 295-303.

[7] M. Choi, A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter, IEEE Transaction on Geoscience and Remote Sensing 44(6) (2006) 1672-1682.

[8] M. Ehlers, Spectral characteristics preserving image fusion based on Fourier domain filtering, Proc. SPIE 5574, Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IV (2004) DOI 10.1117/12.565160.

[9] H. N. Gross, and J. R. Schott, Application of spectral mixture analysis and image fusion techniques for image sharpening, Remote Sens. Environ. 63(2) (1998) 85-94.

[10] P. O. Hoyer, Non-negative sparse coding, in Proc. 12th IEEE Workshop NNSP, (2002) 557-565.

[11] P. O. Hoyer, Non-negative Matrix Factorization with Sparseness Constraints, The Journal of Machine Learning Research 5 (2004) 1457-1469

[12] https://engineering.purdue.edu/ biehl/MultiSpec/hyperspectral.html.

[13] R. L. King, and J. Wang, A wavelet based algorithm for pan sharpening landsat 7 imagery, Geoscience and Remote Sensing Symposium 2 (2001) 849-851.

[14] Y. Ling, M. Ehlers, E. L. Usery, and M. Madden, FFT-enhanced IHS transform method for fusing high-resolution satellite images, ISPRS Journal of Photogrammetry and Remote Sensing 61(6) (2007) 381-392.

[15] U. Kumar, C. Mukhopadhyay, T. V. Ramachandra, Pixel based fusion using IKONOS imagery, International Journal of Recent Trends in Engineering 1(1) (2009) 173-177.

[16] S. Li, and B. Yang. A new pan-Sharpening method using a compressed sensing technique, IEEE Transaction on Geoscience and Remote Sensing 49(2) (2011) 738-746.

[17] S. Li, J. T. Kwok, and Y. Wang, Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images, Information Fusion 3(1) (2002) 17-23.

[18] D. Lee and H. Seung, Learning the parts of objects by non-negative matrix factorization, Nature 401(6755) (1999) 788-791.

[19] T. Lillesand, and R. Kiefer, Remote sensing and image interpretation, 4th ed. New York: Wiley (2000).

[20] C. K. Munechika, J. S. Warnick, C. Salvaggio, and J. R. Schott, Resolution enhancement of multispectral image data to improve classification accuracy, 59(1) (1993) 67-72.

[21] J. M. P. Nascimento, and J. M. B. Dias, Vertex component analysis: a fast algorithm to unmix hyperspectral data, IEEE Transaction on Geoscience and Remote Sensing 43(4) (2005) 898-910.

[22] P. S. Pradhan, R. L. King, N. H. Younan, and D. W. Holcomb, Estimation of the number of decomposition levels for a wavelet-based multiresolution multisensor image fusion, IEEE Transaction on Geoscience and Remote Sensing 44(12) (2006) 3674-3686.

[23] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, Hyperspectral unmixing via $L_{1/2}$ sparsity-constrained nonnegative matrix factorization, IEEE Transaction on Geoscience and Remote Sensing 49(11) (2011) 4282-4297.

[24] G. D. Robinson, H. N. Gross, and J. R. Schott, Evaluation of two applications of spectral mixing models to image fusion, Remote Sens. Environ. 71(3) (2000) 272C281.

[25] G. A. Shaw, and H. K. Burke, Spectral imaging for remote sensing, Lincoln Lab 14(1) (2003) 3-28.

[26] J. B. Tenenbaum, V. Silva, and J. C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290(5500) (2000) 2319-2323.

[27] T.-M. Tu, P. S. Huang, C.-L. Hung, and C.-P. Chang, A fast intensityChueCsaturation fusion technique with spectral adjustment for IKONOS imagery, IEEE Geoscience and Remote Sensing Letters 1(4) (2004) 309-312.

[28] N. Yokoya, T. Yairi, and A. Iwasaki, Coupled non-negative matrix factorization (CNMF) for hyperspectral and multispectral data fusion: application to pasture classification, Geoscience and Remote Sensing symposium (IGARSS) (2011) 1779-1782.

[29] N. Yokoya, T. Yairi, and A. Iwasaki, Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion, IEEE Transaction on Geoscience and Remote Sensing 50(2) (2012) 528-537.

[30] Z. Zhang, and Z. Shi, Nonnegative matrix factorization-based hyperspectral and panchromatic image fusion, Neural Computing and Applications (2012) DOI 10.1007/s00521-012-1014-2.

[31] Z. Zhang, Z. Shi, and Z. An, Hyperspectral and panchromatic image fusion using unmixing-based constrained nonnegative matrix factorization, Optik-International Journal for Light and Electron Optics (2012), http://dx.doi.org/10.1016/j.ijleo.2012.04.022.