Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion

Naoto Yokoya, *Student Member, IEEE*, Takehisa Yairi, and Akira Iwasaki

Abstract—**Coupled nonnegative matrix factorization (CNMF) unmixing is proposed for the fusion of low-spatial-resolution hyperspectral and high-spatial-resolution multispectral data to produce fused data with high spatial and spectral resolutions. Both hyperspectral and multispectral data are alternately unmixed into endmember and abundance matrices by the CNMF algorithm based on a linear spectral mixture model. Sensor observation models that relate the two data are built into the initialization matrix of each NMF unmixing procedure. This algorithm is physically straightforward and easy to implement owing to its simple update rules. Simulations with various image data sets demonstrate that the CNMF algorithm can produce high-quality fused data both in terms of spatial and spectral domains, which contributes to the accurate identification and classification of materials observed at a high spatial resolution.**

Index Terms—**Data fusion, nonnegative matrix factorization, unmixing.**

I. INTRODUCTION

HYPERSPECTRAL imaging sensors collect about 200 spectral band images in the visible and infrared wavelength regions (400–2500 nm). Owing to its high spectral resolution, hyperspectral data are useful for the accurate detection and identification of minerals, vegetation, and man-made materials. The spatial resolution of hyperspectral sensors is often lower than that of multispectral sensors with a low spectral resolution. The fusion of hyperspectral and multispectral data has a possibility to produce fused data with high spatial and spectral resolutions, which contributes to the accurate identification and classification of an area observed at a fine spatial resolution.

There are multiple studies on sharpening algorithms for multispectral data [1]–[10]. Many of these algorithms are designed to fuse multispectral data with a high-spatial-resolution panchromatic image, commonly called pan sharpening. In 2006, the data fusion committee of the IEEE Geoscience and Remote Sensing Society held a public competition for pan sharpening algorithms [9]. The multiresolution methods based on an undecimated discrete wavelet transform or Laplacian pyramid [6] showed the best results. Since the enhancement of the spatial resolution was often limited to the first component, the intensity component, or the low-pass component, a certain amount of spectral distortion occurred [9], [10].

The first algorithm proposed for hyperspectral and multispectral data fusion was a wavelet-based technique that inherited the pan sharpening algorithm [11], [12]. However, its performance highly depended on the spectral resampling method, which caused difficulty in enhancing the spatial resolutions of all hyperspectral band images. A maximum a posteriori (MAP) estimation method was developed to enhance the spatial resolution of hyperspectral data using higher spatial resolution data from an auxiliary sensor [13]–[16]. This approach used a stochastic mixing model (SMM), which estimates the underlying spectral scene characteristics, to develop a cost function that optimizes the estimated hyperspectral data relative to the observed hyperspectral and multispectral data. In the actual implementation, low-spatial-resolution hyperspectral data were processed after principal component (PC) transform [16]. The MAP/SMM method showed a better performance than those based on least-squares estimation [1], [7] and PC substitution [3]. Although the MAP/SMM method with wavelet transforms demonstrated a high noise resistance [17], the experiments were limited to enhancing the low-spatial-resolution hyperspectral band images only in the multispectral wavelength regions.

Another approach for hyperspectral resolution enhancement uses spectral mixture analysis [18], [19]. In this approach, low-spatial-resolution hyperspectral data are unmixed into the endmember spectra and abundances. Next, the abundance maps are fused with high-spatial-resolution panchromatic data using constrained optimization techniques. The results are limited to the synthetic data where the endmembers are known *a priori*. Although this approach did not focus on the estimation of high-spatial-resolution hyperspectral data, the idea of using unmixing for data fusion is physically reasonable and effective for hyperspectral and multispectral data fusion.

In recent decades, many hyperspectral unmixing techniques based on a linear spectral mixture model have been developed [20]–[30]. Unmixing mainly consists of two steps: extracting endmember spectra and calculating their abundance maps. Convex-geometry-based approaches, which assume the presence of at least one pure pixel for each endmember, are commonly used for endmember extraction [24]–[26]. Over the last decade, nonnegative matrix factorization (NMF) [31], [32] has emerged as a useful unmixing method [27]–[30]. Factorizing a nonnegative hyperspectral data matrix into two nonnegative matrices, this method can identify the endmember spectra and can estimate the corresponding abundances simultaneously without pure pixel assumption.

In this paper, we propose a coupled NMF (CNMF) algorithm for hyperspectral and multispectral data fusion based on unsupervised unmixing. Low-spatial-resolution hyperspectral

and high-spatial-resolution multispectral data are alternately unmixed by NMF, taking into account sensor observation models. By combining the hyperspectral endmember matrix and the high-spatial-resolution abundance matrix obtained from multispectral data, high-spatial-resolution hyperspectral data can be generated. Both spatial and spectral reconstruction qualities are evaluated by three simulations using synthetic data sets generated from real airborne hyperspectral data. Comparison results with the MAP/SMM method are also demonstrated. CNMF is applicable to a visible-near-infrared radiometer data set of hyperspectral imager suite (HISUI) [33], which is the Japanese next-generation spaceborne radiometer composed of a hyperspectral radiometer with 30-m ground sampling distance (GSD) and a multispectral radiometer with 5-m GSD. HISUI will provide both hyperspectral and multispectral data obtained over the same region with identical atmospheric and illumination conditions.

This paper is organized as follows. Section II describes the CNMF algorithm for the hyperspectral and multispectral data fusion. Section III presents the characteristics of the simulation data sets and the quantitative criteria for evaluating the performance of the fused data. Experimental results and discussion are presented in Section IV, and the conclusion is given in Section V.

II. CNMF UNMIXING FOR DATA FUSION

The aim of hyperspectral and multispectral data fusion is to estimate unobservable high-spatial-resolution hyperspectral data ($\mathbf{Z} \in \mathbb{R}^{\lambda_h \times L_m}$) from observable low-spatial-resolution hyperspectral data ($\mathbf{X} \in \mathbb{R}^{\lambda_h \times L_h}$) and high-spatial-resolution multispectral data ($\mathbf{Y} \in \mathbb{R}^{\lambda_m \times L_m}$). λ_h and λ_m denote the number of spectral channels of hyperspectral and multispectral sensors, respectively. L_h and L_m denote the number of pixels of hyperspectral and multispectral images, respectively. All data are expressed in a matrix form, with each column vector representing a spectrum at each pixel. $\lambda_h > \lambda_m$ and $L_h < L_m$ are satisfied by the tradeoff between spectral and spatial resolutions of two sensors. We assume that the observed two data are obtained under the same atmospheric and illumination conditions and are geometrically coregistered with radiometric correction.

A. Sensor Observation Model

The spatial domain of the low-spatial-resolution hyperspectral data is degraded from that of the multispectral data. On the other hand, the multispectral data is a spectrally degraded form of the high-spatial-resolution hyperspectral data. Therefore, $\mathbf{X}$ and $\mathbf{Y}$ are modeled as

$$\mathbf{X} = \mathbf{Z}\mathbf{S} + \mathbf{E}_s \tag{1}$$

$$\mathbf{Y} = \mathbf{R}\mathbf{Z} + \mathbf{E}_r. \tag{2}$$

Here, $\mathbf{S} \in \mathbb{R}^{L_m \times L_h}$ is the spatial spread transform matrix, with each column vector $\{\mathbf{s}_{k'}\}_{k'=1}^{L_h} \in \mathbb{R}^{L_m}$ representing the transform of the point spread function (PSF) from the multispectral image to the hyperspectral k'th pixel value. Each

PSF is assumed to be normalized, i.e., $\sum_{k=1}^{L_m} s_{kk'} = 1$. $\mathbf{R} \in \mathbb{R}^{\lambda_m \times \lambda_h}$ is the spectral response transform matrix, with each row vector $\{\mathbf{r}_i\}_{i=1}^{\lambda_m} \in \mathbb{R}^{\lambda_h}$ representing the transform of the spectral response function from the hyperspectral sensor to the multispectral ith band detector. $\mathbf{S}$ and $\mathbf{R}$ are sparse matrices composed of nonnegative components. $\mathbf{E}_s$ and $\mathbf{E}_r$ are the residuals. In the simulation of this paper, $\mathbf{S}$ and $\mathbf{R}$ are given.

When applied to real data, $\mathbf{S}$ is determined by the image registration and estimation of the PSF. $\mathbf{R}$ is derived by radiometric calibration to obtain spectral response functions.

B. Linear Spectral Mixture Model

A linear spectral mixture model is commonly used for unmixing problems owing to its physical effectiveness and mathematical simplicity. A spectrum at each pixel is assumed to be a linear combination of several endmember spectra. Therefore, $\mathbf{Z}$ is formulated as

$$\mathbf{Z} = \mathbf{W}\mathbf{H} + \mathbf{N} \tag{3}$$

where $\mathbf{W} \in \mathbb{R}^{\lambda_h \times D}$ is the spectral signature matrix, with each column vector $\{\mathbf{w}_j\}_{j=1}^{D} \in \mathbb{R}^{\lambda_h}$ representing the endmember spectrum and D being the number of endmembers. $\mathbf{H} \in \mathbb{R}^{D \times L_m}$ is the abundance matrix, with each column vector $\{\mathbf{h}_k\}_{k=1}^{L_m} \in \mathbb{R}^{D}$ denoting the abundance fractions of all endmembers at the pixel, and $\mathbf{N} \in \mathbb{R}^{\lambda_h \times L_m}$ is the residual. The endmember spectra and abundances are nonnegative: $\mathbf{W} \geq 0$ and $\mathbf{H} \geq 0$. In addition, the sum of the abundances for each pixel can be assumed to be unity, i.e., $\sum_{j=1}^{D} h_{jk} = 1 (k = 1, 2, \ldots, L_m)$. When we deal with radiance data, the spectral signatures vary in amplitude owing to the illumination effect caused by the surface topography, buildings, and vegetation. Therefore, with the abundance sum-to-one constraint, the endmember matrix contains several shading endmembers.

By substituting (3) into (1) and (2), $\mathbf{X}$ and $\mathbf{Y}$ can be approximated as

$$\mathbf{X} \approx \mathbf{W}\mathbf{H}_h \tag{4}$$

$$\mathbf{Y} \approx \mathbf{W}_m\mathbf{H}. \tag{5}$$

Here, we define the spatially degraded abundance matrix $\mathbf{H}_h \in \mathbb{R}^{D \times L_h}$ and the spectrally degraded endmember matrix $\mathbf{W}_m \in \mathbb{R}^{\lambda_m \times D}$ given by

$$\mathbf{H}_h \approx \mathbf{H}\mathbf{S} \tag{6}$$

$$\mathbf{W}_m \approx \mathbf{R}\mathbf{W}. \tag{7}$$

Equations (4) and (5) appear as the approximated forms of the linear spectral mixture models degraded in the spatial and spectral domains, respectively. Since the sum of abundances at each pixel in $\mathbf{H}$ is unity and the PSF in $\mathbf{S}$ is normalized, the sum of abundances at each pixel in $\mathbf{H}_h$ is also unity. In addition, owing to the nonnegative characteristics of $\mathbf{S}$, $\mathbf{R}$, $\mathbf{W}$, and $\mathbf{H}$, all components of $\mathbf{H}_h$ and $\mathbf{W}_m$ are also nonnegative. When $\mathbf{S}$ and $\mathbf{R}$ are accurately obtained, the approximations of (6) and (7) become exact.
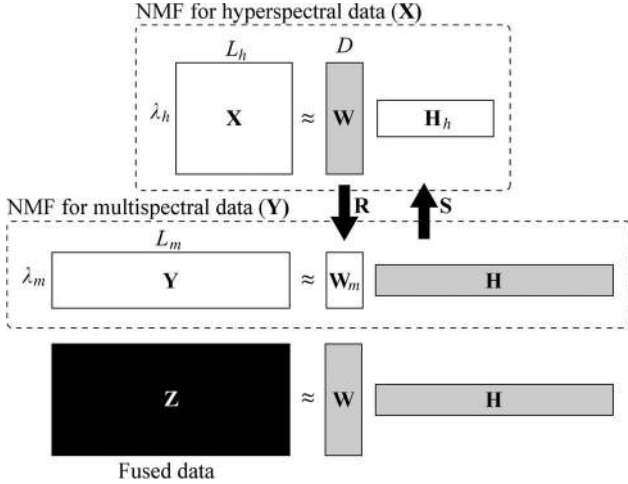
Fig. 1. Illustration of CNMF unmixing for hyperspectral and multispectral data fusion.

C. CNMF Unmixing

Hyperspectral and multispectral data fusion based on unmixing is achieved by the estimation of the high-spectral-resolution endmember spectra and the high-spatial-resolution abundance maps from the two data. The CNMF alternately unmixes $\mathbf{X}$ and $\mathbf{Y}$ by NMF to estimate $\mathbf{W}$ and $\mathbf{H}$, with the constraints (6) and (7) used for the initialization steps. Fig. 1 shows the simplified illustration of the CNMF. Other unmixing and endmember detection algorithms are also applicable to these two unmixing procedures. However, in order to minimize the residual errors in the linear spectral mixture models while considering sensor properties, NMF is straightforward to formulate and easy to implement.

NMF attempts to decompose a nonnegative data matrix into a product of nonnegative matrices [31], [32]. In the case of applying it to a hyperspectral unmixing problem, a squared Frobenius norm of a residual matrix in a linear spectral mixture model is commonly used for a cost function. For the NMF unmixings of $\mathbf{X}$ and $\mathbf{Y}$, these cost functions are defined by $\|\mathbf{X} - \mathbf{W}\mathbf{H}_h\|_F^2$ and $\|\mathbf{Y} - \mathbf{W}_m\mathbf{H}\|_F^2$, respectively, where $\|\cdot\|_F$ denotes the Frobenius norm. Some researchers showed that additional physical constraints to the cost function as penalty terms were effective for hyperspectral unmixing [27]–[31]. However, since the additional penalty functions did not markedly affect the data fusion performance in our experiment, we adopt the simple cost function. To minimize it, Lee and Seung developed multiplicative update rules that are guaranteed to converge to local optima under the nonnegativity constraints of two factorized matrices [31], [32]. We use the multiplicative update rules for the NMF unmixings of $\mathbf{X}$ and $\mathbf{Y}$, which are given as

$$\mathbf{W} \leftarrow \mathbf{W} .* \left(\mathbf{X}\mathbf{H}_h^T\right) ./ \left(\mathbf{W}\mathbf{H}_h\mathbf{H}_h^T\right) \tag{8}$$

$$\mathbf{H}_h \leftarrow \mathbf{H}_h .* \left(\mathbf{W}^T\mathbf{X}\right) ./ \left(\mathbf{W}^T\mathbf{W}\mathbf{H}_h\right) \tag{9}$$

$$\mathbf{W}_m \leftarrow \mathbf{W}_m .* \left(\mathbf{Y}\mathbf{H}^T\right) ./ \left(\mathbf{W}_m\mathbf{H}\mathbf{H}^T\right) \tag{10}$$

$$\mathbf{H} \leftarrow \mathbf{H} .* \left(\mathbf{W}_m^T\mathbf{Y}\right) ./ \left(\mathbf{W}_m^T\mathbf{W}_m\mathbf{H}\right) \tag{11}$$

where $(\cdot)^T$ denotes the transposition of the matrix and $.*$ and $./$ denote elementwise multiplication and division, respectively. To satisfy the abundance sum-to-one constraint, we adopted a method given in [35].

The CNMF algorithm starts from NMF unmixing for the low-spatial-resolution hyperspectral data owing to its spectral advantage. As the initialization phase, we set the number of endmembers D and calculate the initial endmember matrix $\mathbf{W}$ by vertex component analysis (VCA) [26], which is one of the most advanced convex-geometry-based endmember extraction methods with the pure pixel assumption. $\mathbf{H}_h$ is set as the constant value $1/D$ and is updated by (9) until convergence with $\mathbf{W}$ is fixed. As the optimization phase, both $\mathbf{W}$ and $\mathbf{H}_h$ are alternately updated by (8) and (9) until the next convergence. The subsequent round of NMF unmixing for $\mathbf{X}$ differs from the first round described previously only by the initialization phase. $\mathbf{H}_h$ is initialized by (6), and $\mathbf{W}$ is updated by (8) until convergence with $\mathbf{H}_h$ is fixed to inherit the reliable information of abundance maps obtained from multispectral data.

As an alternate step, we apply NMF unmixing to the high-spatial-resolution multispectral data. As the initialization phase, $\mathbf{W}_m$ is set by (7). $\mathbf{H}$ is initialized as the constant value $1/D$ and is updated by (11) until convergence with $\mathbf{W}_m$ is fixed. This process is important in inheriting the reliable information of endmember spectra obtained from hyperspectral data. As the optimization phase, both $\mathbf{W}_m$ and $\mathbf{H}$ are alternately updated by (10) and (11) until the next convergence.

The two NMF unmixing steps are repeated alternately until convergence. The CNMF alternately takes advantages of the spectral information of the low-spatial-resolution hyperspectral data and the spatial information of the multispectral data to find an initialization that results in better local optima. Finally, we can produce the high-spatial-resolution hyperspectral data by multiplying $\mathbf{W}$ with $\mathbf{H}$. We refer to the alternate NMF unmixings as the outer loop and the iterative update in each NMF as the inner loop. As a convergence condition, we use the condition that the change ratio of cost function C achieves a value below a given threshold ε

$$\left|\frac{C^l - C^{l+1}}{C^l}\right| \leq \varepsilon \tag{12}$$

where l is an index of iteration. For a practical utility, the condition that the number of iterations exceeds a predefined maximum number of iterations is added to the stopping criteria together with the convergence condition. The maximum number of iterations is set as different values between the inner and outer loops, referred to as I_{in} and I_{out}, respectively.

I_{in} is a parameter that determines the constraint strength for the sensor properties given by (6) and (7). When we set I_{in} as one and I_{out} as a large value and when there is no update in the initialization phase after the first NMF unmixings for the two data, the subsequent optimization procedures can be simplified as the iterations of (6), (9), (8), (7), (10), and (11) until convergence. In this case, the CNMF algorithm appears to be a projected gradient NMF [36], with the projection steps being (6) and (7), which is mathematically clear and simple to implement. This approach uses tight constraints on the sensor

properties. In contrast, when I_{in} is set as a large value, the CNMF uses loose constraints for the initialization of the NMF unmixing to find better local optima. The CNMF algorithm for hyperspectral and multispectral data fusion is summarized as follows.

Algorithm: CNMF unmixing for hyperspectral and multispectral data fusion
Input: Hyperspectral data $\mathbf{X} \in \mathbb{R}^{\lambda_h \times L_h}$ and multispectral data $\mathbf{Y} \in \mathbb{R}^{\lambda_m \times L_m}$.
Output: Two matrices $\mathbf{W} \in \mathbb{R}^{\lambda_h \times D}$ and $\mathbf{H} \in \mathbb{R}^{D \times L_m}$.
Step 1. First NMF of $\mathbf{X}$
    1a) Initialize $\mathbf{W}$ by VCA, and update $\mathbf{H}_h$ by (9), with $\mathbf{W}$ fixed.
    1b) Optimize $\mathbf{W}$ and $\mathbf{H}_h$ by (8) and (9).
Step 2. NMF of $\mathbf{Y}$
    2a) Initialize $\mathbf{W}_m$ by (7), and update $\mathbf{H}$ by (11), with $\mathbf{W}_m$ fixed.
    2b) Optimize $\mathbf{W}_m$ and $\mathbf{H}$ by (10) and (11).
Step 3. Subsequent NMF of $\mathbf{X}$
    3a) Initialize $\mathbf{H}_h$ by (6), and update $\mathbf{W}$ by (8), with $\mathbf{H}_h$ fixed.
    3b) Optimize $\mathbf{W}$ and $\mathbf{H}_h$ by (8) and (9).
Step 4. Repeat **Steps 2** and **3**.

III. DATA AND EVALUATION

A. Test Data

The proposed hyperspectral and multispectral data fusion technique is applied to three synthetic data sets generated from real airborne hyperspectral data. The first image was taken over Indian Pine by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) [37] sensor in 1996, with 224 spectral bands in the 400–2500-nm region. The second image was also collected by AVIRIS, taken over Cuprite, NV, in 1997. The third image was taken over Washington DC by the Hyperspectral Digital Imagery Collection Experiment (HYDICE) [38] sensor in 1995, with 191 spectral bands in the 400–2500-nm region. We selected these various scenes to evaluate the general effectiveness of CNMF for spectral divergence.

We selected 240×240 pixel size images for all three test data and generated the multispectral and low-spatial-resolution hyperspectral data by downsampling the original hyperspectral data in the spectral and spatial domains, respectively. The multispectral data were produced with uniform spectral response functions corresponding to Landsat TM bands 1–5 and 7, which cover the 450–520-, 520–600-, 630–690-, 760–900-, 1550–1750-, and 2080–2350-nm regions, respectively [15]. The low-spatial-resolution hyperspectral data were generated by a Gaussian PSF with full-width at half maximum, corresponding to six pixels in the original high-spatial-resolution hyperspectral image, which results in a sixfold difference in spatial resolution between two sensors. Therefore, $\mathbf{R}$ and $\mathbf{S}$ are given as sparse matrices, with each row vector corresponding to a uniform spectral response function and with each column vector representing a Gaussian PSF, respectively. In addition, Gaussian noise was added to the two data, supposing that the signal-to-noise ratios (SNRs) of multispectral and hyperspec-

tral sensors are 200 and 300, respectively. The larger spatial-resolution difference and the smaller SNR make the problem of data fusion more difficult. We determined the spatial-resolution difference and SNR of two sensors by considering the specification of HISUI [33]. In this simulation, we assume that two data are obtained under completely identical atmospheric and illumination conditions and are geometrically coregistered.

B. Performance Evaluation

The performance of hyperspectral and multispectral data fusion was evaluated by comparing the estimated high-spatial-resolution hyperspectral data with the original data from two viewpoints: the spatial reconstruction quality of each spectral band image and the spectral reconstruction quality of each spectrum at a single pixel. To evaluate the spatial reconstruction quality, we adopted the peak SNR (PSNR), which is easily defined via the mean square error (MSE). The MSE of the ith spectral band image is defined as

$$\text{MSE}_i = \frac{1}{N} \sum_{k=1}^{N} (\mathbf{Z} - \mathbf{W}\mathbf{H})_{i,k}^2 \qquad (13)$$

where the index (i, k) indicates the kth pixel in the ith band. The PSNR of the ith band is defined as

$$\text{PSNR}_i = 10 \cdot \log_{10} \left(\frac{\text{MAX}_i^2}{\text{MSE}_i} \right) \qquad (14)$$

where MAX_i is the maximum pixel value in the ith band image. A larger PSNR value indicates a higher quality spatial reconstruction. To evaluate the spectral reconstruction quality, we used the spectral angle in the λ-dimensional space between the estimated and actual spectra. A smaller angle indicates a higher quality spectral reconstruction. We refer to this angle as the spectral angle error (SAE).

C. MAP/SMM Method

The MAP/SMM method is one of the most advanced data fusion techniques that can improve the spatial resolution of all hyperspectral band images using high-spatial-resolution multispectral data. The average spectrum, covariance matrix, and abundance map of each endmember are estimated by the SMM. Next, by maximizing the conditional probability density function of $\mathbf{Z}$ given by $\mathbf{X}$ and $\mathbf{Y}$, the high-spatial-resolution hyperspectral data can be calculated. The details of the process are explained, and the MATLAB code is introduced in [16]. In the SMM and MAP processes, the low-spatial-resolution hyperspectral data are transformed into the PCs, and the low-rank PCs are mainly processed, which differs greatly from the proposed method. We use the first six PCs, which have a variance of more than 99.9% for an entire data cube for all test images. Several parameters, such as the number of endmembers, the allowable number of components for a mixed class, and the discrete mixture level, are necessary for the SMM. If the number of endmembers is large, the number of sample spectra for a certain endmember decreases. This results

TABLE I
COMPARISON OF AVERAGE PSNR (IN DECIBELS) AND SAE (IN DEGREES) VALUES, AND COMPUTATIONAL COSTS (IN SECONDS)

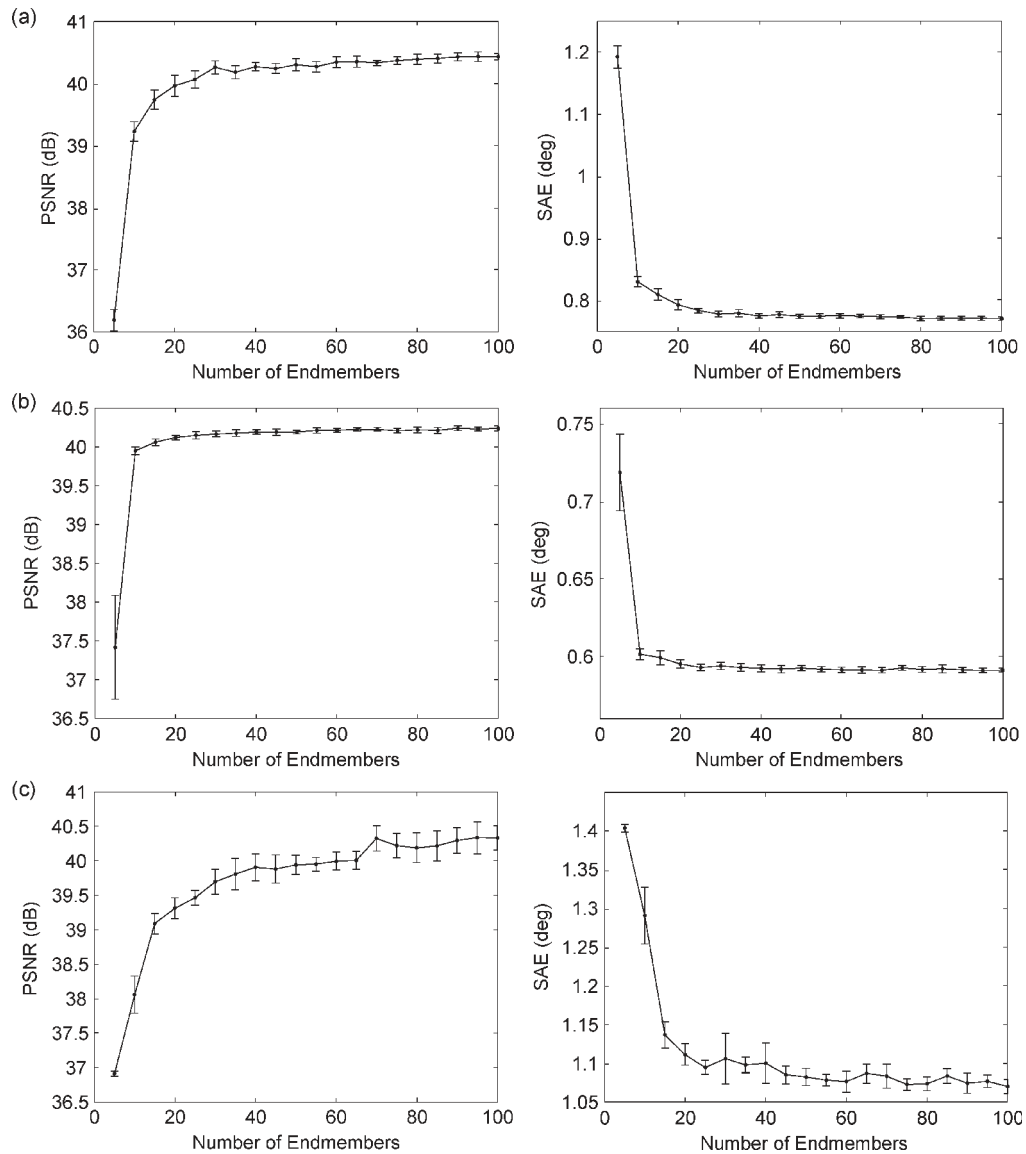|	PSNR (dB)			SAE (deg)			Time (sec)		
	MAP/ SMM	CNMF		MAP/ SMM	CNMF		MAP/ SMM	CNMF	
I_{in}		1	300		1	300		1	300
Indian Pine	39.09	40.06	40.27	0.8677	0.7841	0.7753	94.00	264.07	137.75
Cuprite	37.95	40.17	40.04	0.6544	0.5996	0.5917	93.97	176.82	135.49
Washington DC	37.75	40.20	39.90	1.1511	1.0940	1.0828	91.84	233.78	126.23

Fig. 2.   Effect of endmember number on CNMF fusion quality for the (a) Indian Pine, (b) Cuprite, and (c) Washington DC data.
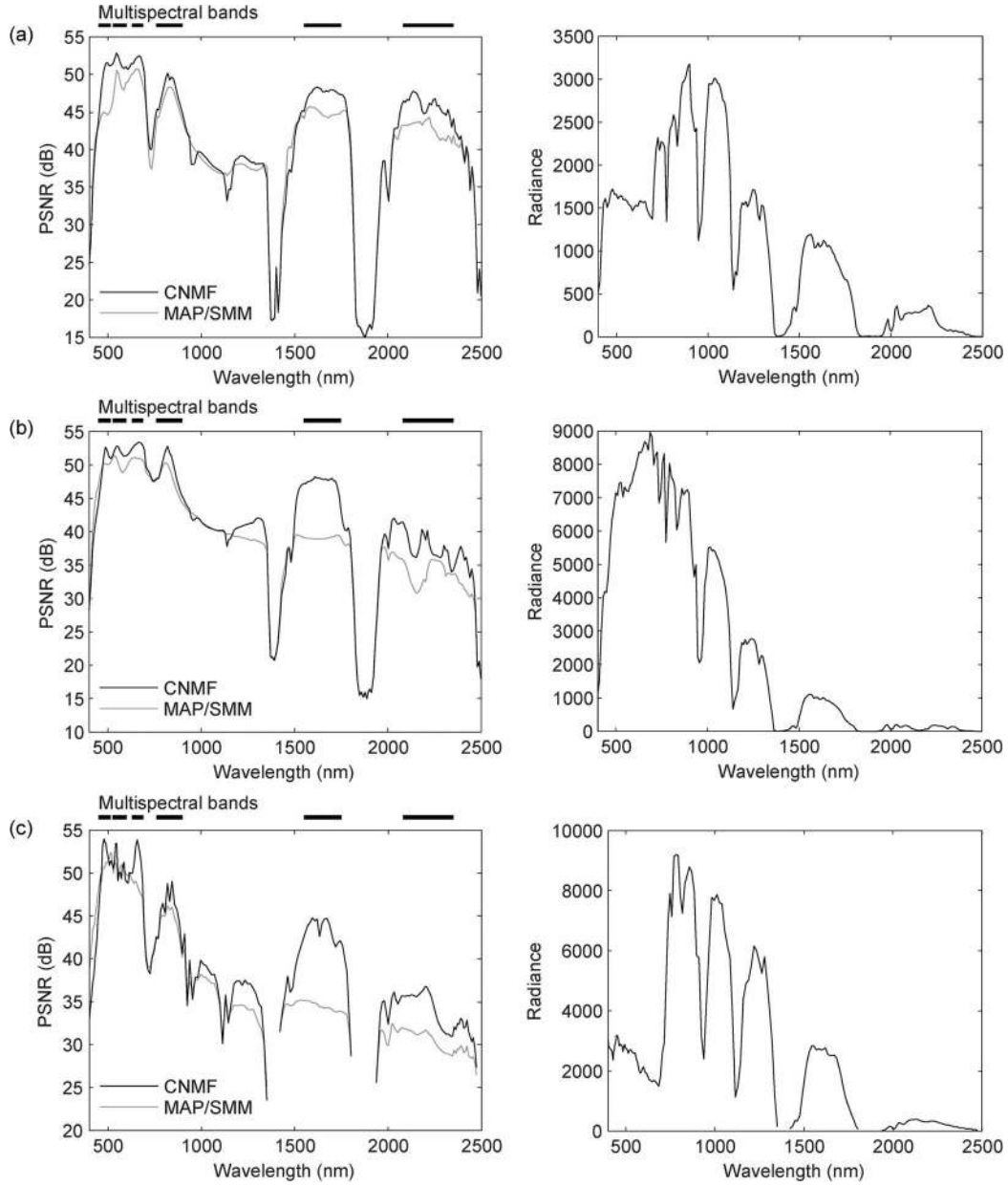
Fig. 3. Comparisons of PSNRs between MAP/SMM and CNMF (left graphs), and typical endmember spectra estimated by CNMF (right graphs) for the (a) Indian Pine, (b) Cuprite, and (c) Washington DC data.

in a nonsingular endmember covariance matrix, which makes the SMM process impossible. Therefore, in this paper, we adopt the SMM with four endmembers and 35 mixture classes that are experimentally set as proper numbers.

IV. RESULTS AND DISCUSSION

We applied the CNMF to the three test data sets. For each data set, we first examined the effects of the maximum number of iterations for inner loops (I_{in}) and the number of endmembers (D). The CNMF method depends on the initial conditions, and VCA produces a slightly different result for each trial. Therefore, we repeated the test ten times for each condition. Next, we evaluated the spatial and spectral performance characteristics of the fused data obtained by the CNMF in detail,

comparing the CNMF with the MAP/SMM using the best results of the ten trials for both algorithms.

A. Constraint Strength

First, the effect of the constraint strength for the sensor properties is examined, setting the number of iterations to two extreme cases. To use tight constraints, the first approach sets $I_{in} = 1$ and $I_{out} = 1000$ without update in the initialization phase after the first NMF unmixings of the two data. The second approach empirically sets $I_{in} = 300$ and $I_{out} = 5$ to use loose constraints. For the first NMF unmixings of the two data, the former adopts the same procedure with the latter. For these two approaches, the thresholds (ε) for the convergence condition and the number of endmembers (D) are set the same at 0.0001
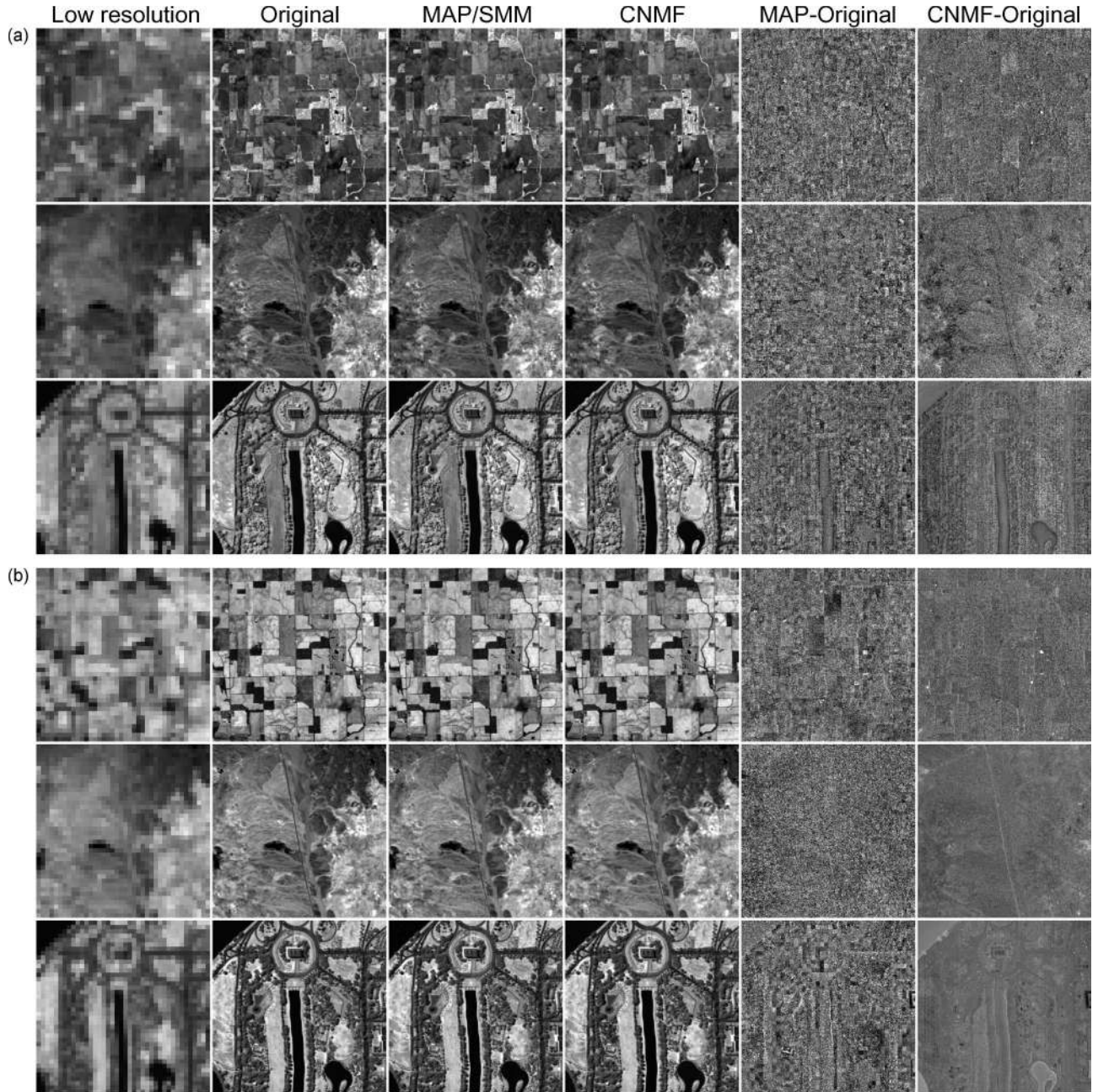
Fig. 4.   Band images in the (a) 815- and (b) 1650-nm regions of low spatial resolution, original high spatial resolution, MAP/SMM, and CNMF, and the difference between hybrid and original images from left to right for three data, i.e., (top) Indian Pine, (middle) Cuprite, and (bottom) Washington DC.

and 40, respectively. The CPU used is Intel(R) Core(TM) i7 CPU 2.80 GHz, with a memory capacity of 16 GB.

Table I shows the average PSNR and SAE values, and the computational cost of the CNMF compared with those of the MAP/SMM. The PSNR and SAE values are the averages of all band images and pixels, respectively. The CNMF methods with two different parameters show similar performance characteristics in both spatial and spectral domains, outperforming the MAP/SMM but taking a longer time. In particular, in the former approach, the computational cost is high owing to the large spatial transform matrix ($\mathbf{S} \in \mathbb{R}^{L_m \times L_h}$). When applied to real data, $\mathbf{S}$ and $\mathbf{R}$ contain errors since it is a challenging issue to estimate the relationship of two sensor properties. Therefore, the latter approach, which uses (6) and (7) fewer times, is

practically effective. Hereafter, we adopt the CNMF algorithm of setting $I_{\text{in}} = 300$ and $I_{\text{out}} = 5$ to examine the effect of D, and we compare its performance in more detail with that of the MAP/SMM.

B. Number of Endmembers

Fig. 2 shows the changes in PSNR and SAE relative to the number of endmembers for three data sets. The error bars show one standard deviation error. As a common tendency, the performance of CNMF improves with an increase in the number of endmembers and finally saturates. This is reasonable because a linear combination of more varying endmember spectra enables a more accurate expression of all pixel spectra in the data until

the endmember variety becomes sufficient. The optimum D can be larger than the actual number of pure materials in the scene because the endmember matrix $\mathbf{W}$ contains shade and mixed material endmember spectra owing to the abundance sum-to-one constraint. The simulation of the Indian Pine data shows a convergence at approximately $D = 40$ in two criteria. In the case of the Cuprite data, both criteria show a faster convergence, and the SAE value is small. In contrast, for the Washington DC data, both criteria converge slowly, and the SAE value is large. These results indicate that the number of endmembers and the final performance depend on scene complexity. The larger the number of endmembers is, the higher is the computational cost. In this paper, we set $D = 40$ as a sufficient number that demonstrates the performance of the CNMF compared with the MAP/SMM in reducing the computational cost.

C. Spatial Performance

The left graphs in Fig. 3 show the PSNR values for all hyperspectral wavelength regions, with the multispectral observation region indicated above the graphs. The right graphs in Fig. 3 show the typical endmember spectra of all of the data sets. For all of the data sets, the CNMF shows comparable or better results than the MAP/SMM in many spectral regions. In most multispectral wavelength regions, the CNMF outperforms the MAP/SMM. This proves that the CNMF can accurately estimate the endmember spectra $\mathbf{W}$ and the abundance fractions $\mathbf{H}$ simultaneously. The difference between the two methods is particularly significant in the short-wave infrared wavelength regions that correspond to Landsat TM bands 5 and 7, which contribute less to the low-rank PCs transformed from the low-spatial-resolution hyperspectral data, owing to a relatively low radiance for all of the data sets. The MAP/SMM is limited in improving the spatial resolution of these wavelength regions because this method enhances only the low-rank PCs. In contrast, since the CNMF is processed in the original data space based on unmixing, it enables the improvement of the spatial resolution in all spectral regions. In the wavelength regions not covered by the multispectral data, the PSNR values for the two methods are relatively low and show no significant difference. This indicates that resolution enhancement is limited without high-spatial-resolution information.

The four column images at the left side of Fig. 4(a) and (b) show the comparisons of band images, respectively, in the 815- and 1650-nm regions between the low-spatial-resolution hyperspectral image, the original high-spatial-resolution hyperspectral image, and the hybrid images estimated by the MAP/SMM and CNMF. We choose these spectral regions as representative regions where the difference in PSNR between the two methods is small and large, respectively. It is difficult to determine the differences between the original image and most of the CNMF and MAP/SMM images with the naked eye. Therefore, we show the radiance difference images between the original images and two hybrid images in the two columns at the right side of Fig. 4. For each datum, the difference images are stretched to the same range. The difference images of the MAP/SMM seem to be random noise images, and the noise variance is larger than that of the CNMF. In the MAP/SMM,
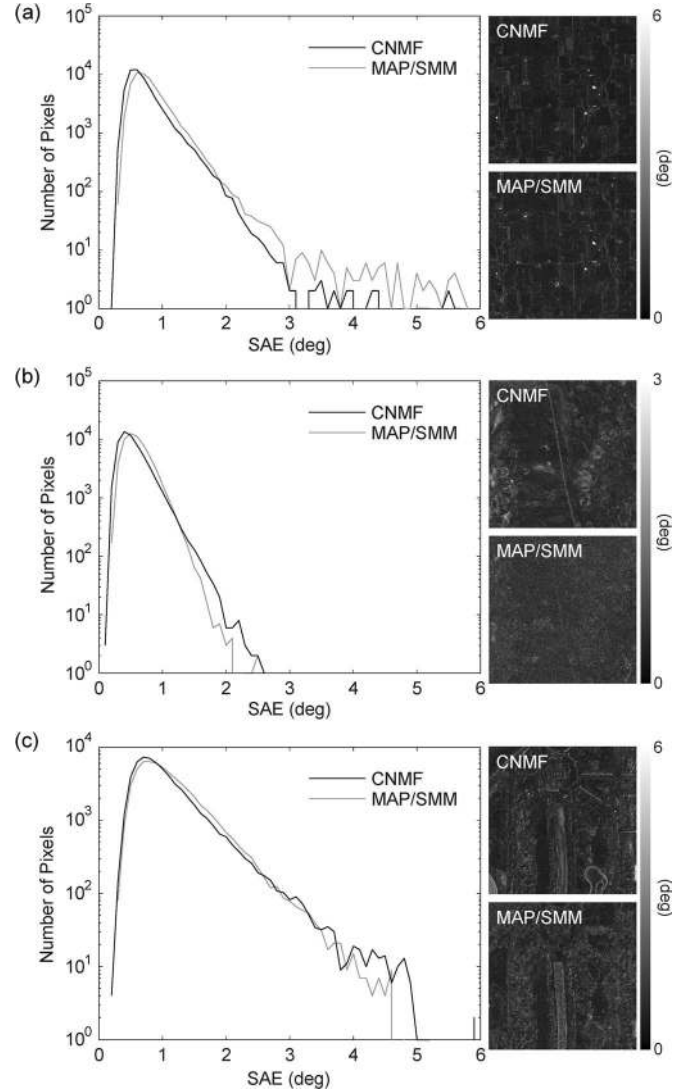

Fig. 5. Histograms of SAE with SAE distribution maps for the (a) Indian Pine, (b) Cuprite, and (c) Washington DC data.

the high-rank PCs of the low-spatial-resolution hyperspectral data, which can be approximated as the random noises, are not processed. Therefore, it is not possible to increase their resolution, and they remain in the fused data and appear in the difference images. In contrast, the image texture is often reflected in the difference images of the CNMF. The accuracy of the estimated endmember spectra affects the reconstruction error of unmixing at each pixel. Therefore, the estimation errors of the CNMF fused data depend on the observed materials, and the image textures appear in the difference images.

D. Spectral Performance

Fig. 5 shows the histograms of SAE [14], [39] and the SAE distribution maps [39]. In these histograms, the CNMF shows comparable or lower spectral errors than the MAP/SMM. As summarized in Table I, while enhancing the spatial resolution of all of the hyperspectral band images, both methods achieve an accuracy of approximately one degree, which allows the accurate identification and classification of an observed area

[39]. The CNMF causes a small spectral distortion owing to the unmixing-based algorithm that is clearly different from pan sharpening algorithms [9], [10]. Since the CNMF enables the increase in the number of endmembers compared with the MAP/SMM, it can deal with spectrally more varied scenes.

The SAE distribution maps demonstrate the property difference between the two fused data. The image textures appear in the CNMF SAE maps owing to the approximation error of NMF unmixing that depends on the radiance value, which is also demonstrated in the spatial performance evaluation. The MAP/SMM SAE maps relatively show a slightly high value but are less affected by the radiance information owing to the probabilistic approach with the Gaussian noise assumption.

V. CONCLUSION

In this paper, we have proposed the CNMF algorithm for hyperspectral and multispectral data fusion. By alternately applying NMF unmixing to low-spatial-resolution hyperspectral and high-spatial-resolution multispectral data, the hyperspectral endmember and high-spatial-resolution abundance matrices are obtained. By combining these two matrices, fused data with both high spatial and spectral resolutions can be obtained. The sensor observation models that relate the two different data are used for the initialization of each NMF step. Therefore, for a practical utility, it is important to determine the sensor properties, such as the PSF and spectral response function. The CNMF is physically straightforward and easy to implement owing to its simple update rules. In addition, it can deal with spectrally varying scenes by setting a large number of endmembers. In the simulations with various image data sets, such as those of vegetation, mineral, and urban with vegetation and water, the CNMF showed comparable or better results than the MAP/SMM, which is one of the most advanced existing algorithms. The high qualities of the CNMF fused data in both the spatial and spectral domains can contribute to the accurate identification and classification of materials observed at a high spatial resolution.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for the many valuable comments and suggestions.

REFERENCES

[1] J. C. Price, "Combining panchromatic and multispectral imagery from dual resolution satellite instruments," *Remote Sens. Environ.*, vol. 21, no. 2, pp. 119–128, Mar. 1987.

[2] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images—II Channel ratio and 'chromacity' transformation techniques," *Remote Sens. Environ.*, vol. 22, no. 3, pp. 343–365, Aug. 1987.

[3] 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," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 265–303, 1991.

[4] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, "The use of intensity-hue-saturation transform for merging SPOT panchromatic and multispectral image data," *Photogramm. Eng. Remote Sens.*, vol. 56, no. 4, pp. 459–467, 1990.

[5] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 1, pp. 49–61, 2000.

[6] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.

[7] R. Nishii, S. Kusanobu, and S. Tanaka, "Enhancement of low spatial resolution image based on high resolution bands," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 5, pp. 1151–1158, Sep. 1996.

[8] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1301–1312, May 2008.

[9] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pan sharpening algorithms: Outcome of the 2006 GRS-S data fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.

[10] Q. Du, N. Younan, R. King, and V. Shah, "On the performance evaluation of pan-sharpening techniques," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 518–522, Oct. 2007.

[11] R. Gomez, A. Jazaeri, and M. Kafatos, "Wavelet-based hyperspectral and multi-spectral image fusion," in *Proc. SPIE*, 2001, vol. 4383, pp. 36–42.

[12] Y. Zhang and M. He, "Multi-spectral and hyperspectral image fusion using 3-D wavelet transform," *J. Electron. (China)*, vol. 24, no. 2, pp. 218–224, Mar. 2007.

[13] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1174–1184, Sep. 2004.

[14] M. T. Eismann and R. C. Hardie, "Application of the stochastic mixing model to hyperspectral resolution enhancement," *IEEE Trans. Geosci.Remote Sens.*, vol. 42, no. 9, pp. 1924–1933, Sep. 2004.

[15] M. T. Eismann and R. C. Hardie, "Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 455–465, Mar. 2005.

[16] M. T. Eismann, "Resolution enhancement of hyperspectral imagery using maximum a posteriori estimation with a stochastic mixing model," Ph.D. dissertation, Univ. Daton, Dayton, OH, May 2004.

[17] Y. Zhang, S. D. Backer, and P. Scheunders, "Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3834–3843, Nov. 2009.

[18] H. N. Gross and J. R. Schott, "Application of spectral mixture analysis and image fusion techniques for image sharpening," *Remote Sens. Environ.*, vol. 63, no. 2, pp. 85–94, Feb. 1998.

[19] G. D. Robinson, H. N. Gross, and J. R. Schott, "Evaluation of two applications of spectral mixing models to image fusion," *Remote Sens. Environ.*, vol. 71, no. 3, pp. 272–281, Mar. 2000.

[20] N. Keshava, "A survey of spectral unmixing algorithms," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 55–78, 2003.

[21] M. O. Smith, P. E. Johnson, and J. B. Adams, "Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis," *J. Geophys. Res.*, vol. 90, no. B2, pp. C797–C804, Oct. 1985.

[22] G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 6, pp. 2706–2717, Nov. 1999.

[23] J. M. P. Nascimento and J. M. B. Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.

[24] J. W. Boardman, "Geometric mixture analysis of imaging spectrometry data," in *Proc. IEEE IGARSS*, Pasadena, CA, 1994, vol. 4, pp. 2369–2371.

[25] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Proc. SPIE*, 1999, vol. 3753, pp. 266–275.

[26] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[27] P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear Algebra Appl.*, vol. 416, no. 1, pp. 29–47, Jul. 2006.

[28] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.

[29] S. Jia and Y. Qian, "Constrained nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 161–173, Jan. 2009.

[30] A. Huck, M. Guillaume, and J. Blanc-Talon, "Minimum dispersion constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2590–2602, Jun. 2010.

[31] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 757–772, Feb. 2011.

[32] D. D. Lee and H. S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.

[33] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Conf. Adv. Neural Inf. Process. Syst.*, 2001, vol. 13, pp. 556–562.

[34] N. Ohgi, A. Iwasaki, T. Kawashima, and H. Inada, "Japanese hyper–multi spectral mission," in *Proc. IGARSS*, Honolulu, HI, Jul. 2010, pp. 3756–3759.

[35] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

[36] C.-J. Lin, "Projected gradient methods for non-negative matrix factorization," Dept. Comput. Sci., Nat. Taiwan Univ., Taipei, Taiwan, Information and Support Services Tech. Rep. ISSTECH-95-013, 2005.

[37] G. Vane, R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen, and W. M. Porter, "The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 44, no. 2/3, pp. 127–143, May/Jun. 1993.

[38] R. Basedow, D. Armer, and M. Anderson, "HYDICE system: Implementation and performance," in *Proc. SPIE*, Jun. 1995, vol. 2480, pp. 258–267.

[39] S. Cook, J. Harsanyi, and V. Faber, "Evaluation of algorithms for compressing hyperspectral data," in *Proc. SPIE*, 2004, vol. 5234, pp. 712–717.

Takehisa Yairi received the M.Sc. and Ph.D. degrees in aerospace engineering from the University of Tokyo (UT), Tokyo, Japan, in 1996 and 1999, respectively.

He is currently an Associate Professor with the Department of Advanced Interdisciplinary Studies, UT. His research interests include machine learning theory and its application to satellite data.

Akira Iwasaki received the M.Sc. degree in aerospace engineering and the Ph.D. degree in engineering from the University of Tokyo, Tokyo, Japan, in 1987 and 1996.

He joined the Electrotechnical Laboratory in 1987 where he engaged in research on space technology and remote sensing system. He is currently a Professor with the University of Tokyo.

Dr. Iwasaki is the Secretary-General of the Remote Sensing Society of Japan.

Naoto Yokoya (S'10) received the B.S. and M.S. degrees in aerospace engineering from the University of Tokyo, Tokyo, Japan, in 2008 and 2010, respectively, where he is working toward the Ph.D. degree in aerospace engineering.

His research interests include hyperspectral data analysis.