Cross-Calibration for Data Fusion of EO-1/Hyperion and Terra/ASTER

Naoto Yokoya, Norimasa Mayumi, and Akira Iwasaki

Abstract—The data fusion of low spatial-resolution hyperspectral and high spatial-resolution multispectral images enables the production of high spatial-resolution hyperspectral data with small spectral distortion. EO-1/Hyperion is the world's first hyperspectral sensor. It was launched in 2001 and has a similar orbit to Terra/ASTER. In this work, we apply hyperspectral and multispectral data fusion to EO-1/Hyperion and Terra/ASTER datasets by the preprocessing of datasets and the onboard cross-calibration of sensor characteristics. The relationship of the spectral response function is determined by convex optimization by comparing hyperspectral and multispectral images over the same spectral range. After accurate image registration, the relationship of the point spread function is obtained by estimating a matrix that acts as Gaussian blur filter between two images. Two pansharpening-based methods and one unmixing-based method are adopted for hyperspectral and multispectral data fusion and their properties are investigated.

Index Terms—Cross-calibration, data fusion, hyperspectral, multispectral.

I. INTRODUCTION

HYPERION [1] and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) [2] data are currently available as real satellite hyperspectral and multispectral datasets. Hyperion is a hyperspectral imager onboard the Earth Observing 1 (EO-1) satellite launched on 21 November 2000. The Hyperion system, consisting of visible near-infrared radiometer (VNIR) and shortwave infrared radiometer (SWIR) subsystems, acquires data in 198 spectral bands with 30 m ground sampling distance. The Hyperion VNIR obtains optical images in 50 bands over a range of 427–925 nm, and the full width at half maximum (FWHM) of the spectral response function is 10.2 nm. Owing to this narrow spectral width, Hyperion can be used for spectral cross-calibration with a multispectral sensor. ASTER is a multispectral imager on the Terra platform, which was launched on 18 December 1999. ASTER consists

of a VNIR, a SWIR, and a thermal infrared radiometer (TIR) with 14 spectral bands. The ASTER VNIR has three spectral channels over a spectral range of 520–860 nm with 15 m ground sampling distance. The Hyperion/VNIR and ASTER/VNIR datasets are good examples of hyperspectral and multispectral datasets with trade-offs between spatial and spectral resolution.

Hyperspectral imaging sensors generally have a larger ground sampling distance than multispectral imaging sensors. The data fusion of low spatial-resolution hyperspectral and high spatial-resolution multispectral images can produce high spatial-resolution hyperspectral data with small spectral distortion [3]–[7]. The fused data enable accurate classification with a fine spatial resolution and thereby enhance the applications of hyperspectral remote sensing, e.g., urban classification and tree-level forest monitoring [8], [9] . There are many studies on pansharpening techniques, which enhance the spatial resolution of multispectral data using panchromatic data [10]–[13]. Compared with pansharpening, hyperspectral and multispectral data fusion is a new technology. Several hyperspectral and multispectral fusion methods have been proposed that use pansharpening techniques, a stochastic method, and unmixing [3]–[7]. Pansharpening based methods such as smoothing filter-based intensity modulation (SFIM) [10] (called the blur transform in [7]) and spectral simulation color normalization (SSCN) have produced effective results in the common spectral regions of two sensors [7]. Since hyperspectral images and the multispectral images in the corresponding band are highly correlated, the spectral distortion of fused data becomes small. This is different from the case of pansharpening, i.e., the fusion of multispectral and panchromatic images. A maximum a posteriori estimation method was developed to enhance the spatial resolution of hyperspectral data using higher spatial-resolution data such as multispectral and panchromatic images [5]. An unmixing-based fusion method, coupled nonnegative matrix factorization (CNMF), was recently proposed to enhance the spatial resolution of all hyperspectral bands [6]. In many studies, data fusion methods have only been evaluated using synthetic datasets generated from original high spatial-resolution hyperspectral data using simple sensor characteristics. For practical use, methods of applying hyperspectral and multispectral data fusion to a real dataset including the preprocessing and estimation of sensor characteristics need to be established.

In the application of real data, the preprocessing of datasets and the onboard cross-calibration of the characteristics (point spread function and spectral response function) of hyperspectral and multispectral sensors are necessary for accurate data fusion. Image registration and radiometric correction are considered as preprocessing techniques. In this work, we apply hyperspectral and multispectral data fusion to Hyperion and ASTER datasets.

A method for the cross-calibration of spatial and spectral characteristics is proposed. SFIM, SSCN, and CNMF are adopted as the fusion methods and their properties are investigated.

This paper is organized as follows. In Section II we describe the SFIM, SSCN, and CNMF data fusion methods. In Section III we present preprocessing and estimation methods to determine the relationships between sensor characteristics in both spatial and spectral domains. In Section IV we describe experimental data and evaluation metric. Experimental results and a discussion are presented in Section V, and the paper is concluded in Section VI.

II. DATA FUSION METHODS

The aim of hyperspectral and multispectral data fusion is to estimate unobservable high spatial-resolution hyperspectral data ($\mathbf{Z} \in \mathbb{R}^{L_h \times N_m}$) from observed low spatial-resolution hyperspectral data ($\mathbf{X} \in \mathbb{R}^{L_h \times N_h}$) and high spatial-resolution multispectral data ($\mathbf{Y} \in \mathbb{R}^{L_m \times N_m}$). L_h and L_m denote the numbers of spectral channels of hyperspectral and multispectral sensors, respectively. N_h and N_m denote the numbers of pixels in hyperspectral and multispectral images, respectively. Owing to the trade-off between the spectral and spatial resolution of the two sensors, $L_h > L_m$ and $N_h < N_m$ are satisfied. All data are expressed in matrix form with each column vector representing the spectrum at each pixel. The two observed datasets are assumed to be obtained under the same atmospheric and illumination conditions, and to be geometrically co-registered with a radiometric correction.

Hyperspectral and multispectral images can be considered as degraded versions of a high spatial-resolution hyperspectral image in the spatial and spectral domains, respectively. When these degradations are assumed to be linear operations, $\mathbf{X}$ and $\mathbf{Y}$ are modeled as

$$\mathbf{X} = \mathbf{ZS} + \mathbf{E}_s, \tag{1}$$
$$\mathbf{Y} = \mathbf{RZ} + \mathbf{E}_r. \tag{2}$$

Here, $\mathbf{S} \in \mathbb{R}^{N_m \times N_h}$ is the point spread transform matrix, which acts as a low-pass filter. Each column vector $\{\mathbf{s}_l\}_{l=1}^{N_h} \in \mathbb{R}^{N_m}$ represents the transform of the point spread function from the multispectral image to the lth pixel value in the hyperspectral image and is assumed to be normalized, i.e., $\sum_{k=1}^{L_m} s_{kl} = 1$. $\mathbf{R} \in \mathbb{R}^{L_m \times L_h}$ is the spectral response transform matrix with each row vector $\{\mathbf{r}_i\}_{i=1}^{L_m} \in \mathbb{R}^{L_h}$ representing the transform of the spectral response function from the hyperspectral sensor to the ith band detector in the multispectral sensor. $\mathbf{E}_s$ and $\mathbf{E}_r$ are residuals. Methods of estimating $\mathbf{S}$ and $\mathbf{R}$ are given in the next section. In this section, we introduce hyperspectral and multispectral data fusion methods based on pan-sharpening and unmixing.

A. Pansharpening-Based Method

Pansharpening is a technique for enhancing the spatial resolution of a multispectral image by using a panchromatic image. It is applicable to hyperspectral and multispectral data fusion over the overlapping spectral bands [7]. In this section, we introduce two pansharpening-based methods for hyperspectral and multispectral data fusion, i.e., SFIM and SSCN, because they

produce good results compared with the commonly used and state-of-the-art pansharpening methods despite their simple implementation [7].

SFIM assumes that the radiance ratio between high spatial-resolution multispectral and original panchromatic images is the same as that between original multispectral and low-pass filtered panchromatic images based on a simplified model for solar radiation and land surface reflection [10]. When applied to hyperspectral and multispectral data fusion, the relationship between multispectral and panchromatic images can be replaced with that between hyperspectral and multispectral images. Therefore, the kth ($k = 1, 2, \ldots, N_m$) pixel value of the high spatial-resolution hyperspectral image in the jth ($j = 1, 2, \ldots, L_h$) band is given by

$$\mathbf{Z}_{j,k} = \frac{\mathbf{Y}_{i,k} \mathbf{X}_{j,l}}{\tilde{\mathbf{Y}}_{i,l}}, \tag{3}$$

where $\tilde{\mathbf{Y}} \in \mathbb{R}^{L_m \times N_h}$ is the low spatial-resolution multispectral image, i is the index of the multispectral band that covers the spectral range of the jth hyperspectral band, and l ($l = 1, 2, \ldots, N_h$) is the pixel index in the low spatial-resolution image corresponding to k ($k = 1, 2, \ldots, N_m$) in the high spatial-resolution image. The low spatial-resolution multispectral image can be given by a low-pass filtered version of the original multispectral image:

$$\tilde{\mathbf{Y}} \approx \mathbf{YS}. \tag{4}$$

Substituting $\mathbf{YS}$ for $\tilde{\mathbf{Y}}$ in (3), in SFIM the kth ($k = 1, 2, \ldots, N_m$) pixel value of the high spatial-resolution hyperspectral image in the jth ($j = 1, 2, \ldots, L_h$) band is calculated as

$$\mathbf{Z}_{j,k} = \frac{\mathbf{Y}_{i,k} \mathbf{X}_{j,l}}{(\mathbf{YS})_{i,l}}. \tag{5}$$

SSCN is an extension of the color normalization (CN) method [4]. It is also based on the assumption of (3), and the low spatial-resolution multispectral image ($\tilde{\mathbf{Y}} \in \mathbb{R}^{L_m \times N_h}$) is assumed to be written as a linear combination of the original hyperspectral images. The spectral response functions of two sensors are taken into account in the coefficients used in the linear combination instead of the simple averaging used in the CN method. The low spatial-resolution multispectral image can be given by

$$\tilde{\mathbf{Y}} \approx \mathbf{RX}. \tag{6}$$

This fundamental assumption of the SSCN method is valid for hyperspectral and multispectral datasets owing to the sufficiently narrow hyperspectral bandwidth. By substituting $\mathbf{RX}$ for $\tilde{\mathbf{Y}}$ in (3), the kth ($k = 1, 2, \ldots, N_m$) pixel value of the high spatial-resolution hyperspectral image in the jth ($j = 1, 2, \ldots, L_h$) band is obtained as

$$\mathbf{Z}_{j,k} = \frac{\mathbf{Y}_{i,k} \mathbf{X}_{j,l}}{(\mathbf{RX})_{i,l}}, \tag{7}$$

B. Unmixing-Based Method

The CNMF method based on linear unmixing was proposed for hyperspectral and multispectral data fusion [6]. This method

gives good results for synthetic hyperspectral and multispectral datasets with a six fold space difference in the ground sampling distance. Over the last decade, nonnegative matrix factorization (NMF) [14], [15] has emerged as a useful unmixing method [16]–[20]. The CNMF method is composed of alternate NMF unmixings of hyperspectral and multispectral images, which are used to extract high spectral-resolution endmember spectra and high spatial-resolution abundance maps.

In a linear spectral mixture model, the spectrum at each pixel is assumed to be a linear combination of several endmember spectra. Therefore, $\mathbf{Z}$ can be formulated as

$$\mathbf{Z} = \mathbf{W}\mathbf{H} + \mathbf{E}, \tag{8}$$

where $\mathbf{W} \in \mathbb{R}^{L_h \times D}$ is the spectral signature matrix with each column vector representing an endmember spectrum, and D is the number of endmembers. $\mathbf{H} \in \mathbb{R}^{D \times N_m}$ is the abundance matrix, with each column vector denoting the abundance fractions of all endmembers at the pixel, and $\mathbf{E} \in \mathbb{R}^{L_h \times N_m}$ is the noise. The endmember spectra and abundances are nonnegative. In addition, the sum of the abundances for each pixel can be assumed to be unity, i.e., $\sum_{j=1}^{D} h_{jk} = 1$ $(k = 1, 2, \ldots, N_m)$.

The spatially degraded abundance matrix $\mathbf{H}_h \in \mathbb{R}^{D \times N_h}$ and the spectrally degraded endmember matrix $\mathbf{W}_m \in \mathbb{R}^{L_m \times D}$ are defined as $\mathbf{H}_h = \mathbf{H}\mathbf{S}$ and $\mathbf{W}_m = \mathbf{R}\mathbf{W}$, respectively. By substituting (8) into (1) and (2), $\mathbf{X}$ and $\mathbf{Y}$ can be approximated as

$$\mathbf{X} \approx \mathbf{W}\mathbf{H}_h, \tag{9}$$

$$\mathbf{Y} \approx \mathbf{W}_m\mathbf{H}. \tag{10}$$

Owing to the nonnegative characteristics of $\mathbf{S}$, $\mathbf{R}$, $\mathbf{W}$, and $\mathbf{H}$, all the components of $\mathbf{H}_h$ and $\mathbf{W}_m$ are also nonnegative. The CNMF algorithm starts from the NMF unmixing of a hyperspectral image ($\mathbf{X}$) in order to use its spectral advantage. Hyperspectral ($\mathbf{X}$) and multispectral ($\mathbf{Y}$) images are alternately unmixed by NMF using $\mathbf{H}_h = \mathbf{H}\mathbf{S}$ and $\mathbf{W}_m = \mathbf{R}\mathbf{W}$ for the initialization of $\mathbf{H}_h$ and $\mathbf{W}_m$, respectively. The high spatial-resolution hyperspectral image is obtained by the multiplication of $\mathbf{W}$ and $\mathbf{H}$. To satisfy the abundance sum-to-one constraint, a method in [21] is used. More details concerning CNMF are given in [6].

III. PREPROCESSING AND CROSS-CALIBRATION OF SENSOR CHARACTERISTICS

A. Spatial Domain

In this section, we present the proposed methods for the image registration of Hyperion and ASTER images and the estimation of the point spread transform matrix ($\mathbf{S}$). In general, the spatial resolution of multispectral sensors is superior to that of hyperspectral sensors. The ground sampling distance of an ASTER VNIR image is 15 m whereas that of a Hyperion image is 30 m. A template-matching technique using correlation coefficient (CC) is applied to find the corresponding coordinates in Hyperion and ASTER images for each pixel [22]. The CCs between the pixel with the maximum value and the neighboring pixels are fitted to a parabola function to calculate the corresponding coordinates with subpixel order. Both the ASTER and Hyperion images undergo warping because of satellite

vibration, the Earth's rotation, elevation parallax, and sensor parallax. Accurate registration of the two images is essential to obtain high-quality fused images [23]. In this study, some contrivances are used in the search for corresponding points. In regions without characteristics such as seas, lakes, and rivers, it is difficult to find corresponding coordinates. Therefore, in such regions, we estimate the corresponding coordinates using affine transformations obtained from successful regions. The coefficients of affine transformations are calculated by the least-squares method using reliable matching coordinates. The ASTER images are resampled and geometrically registered to the Hyperion coordinate.

After the image registration, the point spread transform matrix ($\mathbf{S}$) is estimated by finding the relation between the image sharpness of two low spatial-resolution multispectral images, i.e., $\mathbf{Y}\mathbf{S}$ and $\mathbf{R}\mathbf{X}$. The lth $(l = 1, 2, \ldots, N_h)$ column vector of matrix $\mathbf{S}$ is assumed to be a discrete approximation of a 2D Gaussian function for the lth pixel value in the hyperspectral image. Therefore, matrix $\mathbf{S}$ acts as a Gaussian blur filter for down-sampling the spatial resolution of the multispectral image. For each multispectral band, a Gaussian blur filter is estimated on a grid for a Gaussian blur parameter by maximizing the CC between the two Sobel images of the low spatial-resolution multispectral images obtained from $\mathbf{Y}\mathbf{S}$ and $\mathbf{R}\mathbf{X}$. The Sobel operator is a common edge detection algorithm and is used to calculate the sharpness of an image [24].

$\mathbf{S}$ and $\mathbf{R}$ are necessary for the estimation of $\mathbf{R}$ and $\mathbf{S}$, respectively. We have empirically found that a change in the spectral response transform matrix ($\mathbf{R}$) after the launch does not change the sharpness of the low spatial-resolution multispectral image ($\mathbf{R}\mathbf{X}$). Therefore, the estimation of $\mathbf{S}$ is first conducted using $\mathbf{R}$ obtained from the prelaunch sensor characteristics.

B. Spectral Domain

In this section, the proposed method for estimating the relative spectral response function using convex optimization is presented. Let us consider $\mathbf{X} \in \mathbb{R}^{L_h \times N_h}$ and $\tilde{\mathbf{Y}} \in \mathbb{R}^{L_m \times N_h}$ as hyperspectral and multispectral images with the same spatial resolution, respectively. $\tilde{\mathbf{Y}}$ is obtained as a low-pass filtered multispectral image by (4) using matrix $\mathbf{S}$. When the spectral bandwidths of a hyperspectral sensor are extremely narrow and cover the entire spectral range of a multispectral sensor, the spectral response functions for the multispectral sensor can be approximated as a linear combination of those for the hyperspectral sensor. In this case, the low spatial-resolution multispectral image ($\tilde{\mathbf{Y}}$) can be assumed to be a linear combination of the original hyperspectral images ($\mathbf{X}$) as shown in (6). The ith row vector ($\{\mathbf{r}_{i=1}^{L_m}\} \in \mathbb{R}^{1 \times L_h}$) of matrix $\mathbf{R}$ consist of the coefficients of the linear combination for the ith multispectral band. $\mathbf{R}_{i,j}$ which is the ith row and j^{th} column component of $\mathbf{R}$, denotes the contribution ratio of the jth hyperspectral band to the ith multispectral band. $\mathbf{R}_{i,j}$ is estimated by calculating the degree of mutual overlapping of the spectral response functions between the ith multispectral band and the jth hyperspectral band:

$$\mathbf{R}_{i,j} = \frac{\int f_i(\lambda)g_j(\lambda)d\lambda}{\sum_{j'} \int f_i(\lambda)g_{j'}(\lambda)d\lambda}, \tag{11}$$

where $\{f_i(\lambda)\}_{i=1}^{L_m}$ is the spectral response function of the ith multispectral band and $\{g_j(\lambda)\}_{j=1}^{L_h}$ is that of the jth hyperspectral band, both of which are functions of the wavelength λ. The spectral response functions of satellite optical sensors can be changed after launching [25]. $\mathbf{R}^*$ is used as the prelaunch $\mathbf{R}$, which is obtained from (11) using the prelaunch spectral response functions. When $\mathbf{R}^*$ is used for (6), the approximation of (6) contains large errors because of the difference between the histograms of $\tilde{\mathbf{y}}_i \in \mathbb{R}^{1 \times N_m}$ and $\mathbf{r}_i \mathbf{X}$ $(i = 1, 2, \ldots, L_m)$. For each multispectral band, $\tilde{\mathbf{y}}_i$ is radiometrically modified by histogram matching. Onboard estimation of the spectral response function requires the optimization of $\mathbf{R}$ that satisfies (6) using the observed $\mathbf{X}$ and $\tilde{\mathbf{Y}}$ as well as the prior knowledge of $\mathbf{R}^*$. When we focus on the image of the ith multispectral band ($\tilde{\mathbf{y}}_i \in \mathbb{R}^{1 \times N_m}$), it is given in terms of the spectral response transform vector $\mathbf{r}_i \in \mathbb{R}^{1 \times L_h}$ as

$$\tilde{\mathbf{y}}_i \approx \mathbf{r}_i \mathbf{X}. \tag{12}$$

When we assume that $\mathbf{R}$ is different from $\mathbf{R}^*$ by a ratio of $1 \pm \varepsilon$ ($\varepsilon \geq 0$), the estimation of $\mathbf{R}$ boils down to the following minimization with respect to $\mathbf{r}_i$ $(i = 1, 2, \ldots, L_m)$:

$$\text{minimize} \quad \|\tilde{\mathbf{y}}_i - \mathbf{r}_i \mathbf{X}\|_2^2, \tag{13}$$

$$\text{subject to} \quad |\mathbf{r}_i - \mathbf{r}_i^*| \leq \mathbf{r}_i^* \varepsilon, \quad \mathbf{r}_i^* \geq 0. \tag{14}$$

The constraint given by (14) can be reformulated into a single inequality as

$$\begin{pmatrix} -\mathbf{I} \\ \mathbf{I} \end{pmatrix} \mathbf{r}_i \leq \begin{pmatrix} -\mathbf{r}_i^*(1 - \varepsilon) \\ \mathbf{r}_i^*(1 + \varepsilon) \end{pmatrix}, \tag{15}$$

where $\mathbf{I} \in \mathbb{R}^{L_h \times L_h}$ is the identity matrix. This minimization problem requires convex optimization [26]. We used CVX for convex optimization, which is a MATLAB-based modeling system [27].

IV. Data and Evaluation

Hyperspectral and multispectral data fusion was applied to Hyperion and ASTER images taken over San Francisco on 31 July, 2002. On that date, two sensors were in the same orbit with a 30 min time difference (Hyperion: 10:35 am, ASTER: 11:05 am). We assumed that the observation conditions for these two images were the same. The Hyperion/VNIR data with 30 m ground sampling distance and 50 bands (spectral channels 8–57) and ASTER/VNIR data with 15 m ground sampling distance and there bands were used for the fusion of hyperspectral and multispectral data.

The sensor characteristics (matrices $\mathbf{R}$ and $\mathbf{S}$) are estimated using the original datasets. To evaluate the performances of the three data fusion techniques in Section II, two scenarios are considered.

1) In the first scenario, the original images are spatially degraded and the fused image is compared with the original hyperspectral image.
2) In the second scenario, the quality of data fusion is evaluated by comparing hyperspectral and multispectral images degraded from the fused image with the original images.

In the evaluation of the image-sharpening technique for real data, the first scenario is common because many pansharpening
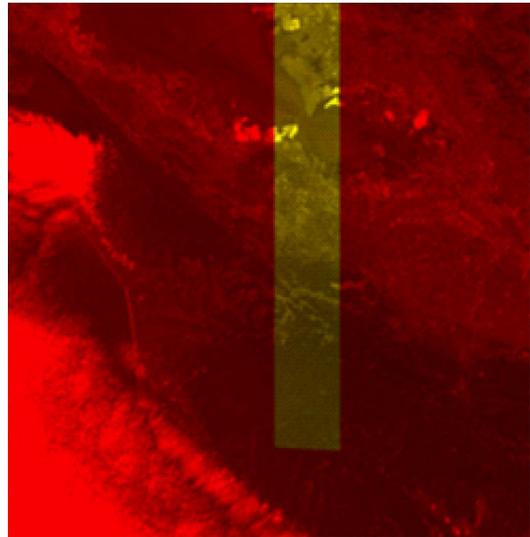

Fig. 1. Registered Hyperion (yellow) and ASTER (red) images.

techniques mainly focus on enhancing the resolution of multispectral images. In this scenario, $\mathbf{R}$ is the same as that estimated from the original datasets and $\mathbf{S}$ is estimated from the spatially degraded datasets. The first scenario is biased toward information in the high spectral-resolution image. There is no guarantee that the fused image will also inherit the properties of the low spectral-resolution and high spatial-resolution images. We regard hyperspectral and multispectral data fusion as not only enhancing the spatial resolution of the hyperspectral image but also enhancing the spectral resolution of the multispectral image. The second scenario is adopted because it is important that the fused image inherits the information in both the hyperspectral and multispectral images. As a result of this evaluation, we can determine the characteristics of the fusion methods.

The root-mean-square error (RMSE) and CC are used to evaluate the spatial reconstruction quality of each band image, and spectral angle error (SAE) is adopted for the spectral evaluation of each pixel. To compare the pansharpening-based methods and the unmixing-based method, the overlapping spectral bands of the two sensors, i.e., Hyperion bands 13–24 for ASTER band 1, Hyperion bands 25–35 for ASTER band 2, and Hyperion bands 36–53 for ASTER band 3, are used to evaluate and characterize their performances.

V. Results and Discussion

A. Sensor Characteristics

The Hyperion image registered to the ASTER coordinates is shown in Fig. 1. The difference between the actual corresponding coordinates and those estimated by affine transformation is shown in Fig. 2. Satellite vibration and sensor parallax can be observed in Fig. 2. The coefficients of the affine transformation between ASTER band 1 and Hyperion are calculated as

$$\begin{pmatrix} x_{\text{ASTER}} \\ y_{\text{ASTER}} \end{pmatrix} = \begin{pmatrix} 1.0059 & 0.0005 \\ 0.0530 & 1.0171 \end{pmatrix} \begin{pmatrix} x_{\text{Hyperion}} \\ y_{\text{Hyperion}} \end{pmatrix} + \begin{pmatrix} 1041.9 \\ -1212.5 \end{pmatrix} \tag{16}$$
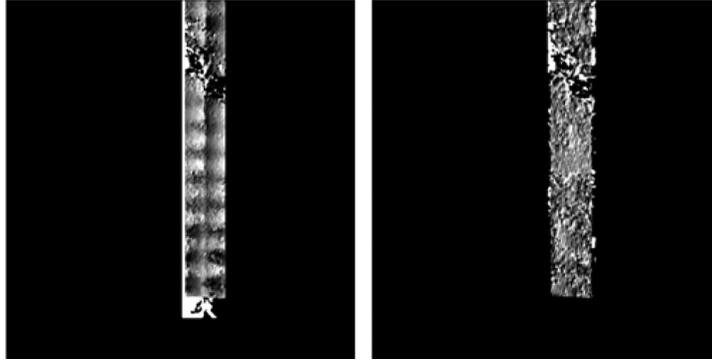
Fig. 2. Residual of coordinates obtained from affine transformation (left-x, right-y).
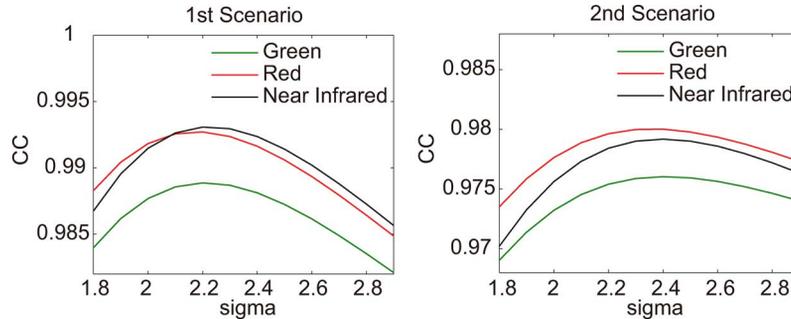

Fig. 3. Correlation coefficient of sharpness between two low spatial-resolution multispectral images generated from Hyperion and ASTER data.
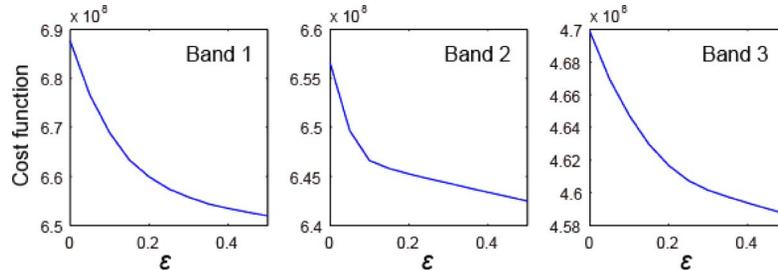

Fig. 4. Effect of relative change in onboard spectral response transform matrix (ε) on cost function.

where $(x_{\text{ASTER}}, y_{\text{ASTER}})$ and $(x_{\text{Hyperion}}, y_{\text{Hyperion}})$ denote the column and row coordinates of the ASTER and Hyperion images, respectively. Fig. 3 shows the CC of the Sobel images between two low spatial-resolution multispectral images generated from Hyperion ($\mathbf{r}_i\mathbf{X}$) and ASTER ($\mathbf{y}_i\mathbf{S}$) ($i = 1, 2, 3$) data. The CC is a function of the FWHM of the Gaussian blur function. Finally, we set the FWHM as 2.2 and 2.4 for the first and second scenarios, respectively. This means that the Hyperion image is more blurred than the official difference between ground sampling distances of Hyperion and ASTER, i.e., 2.

Fig. 4 shows the effect of the relative change in the onboard spectral response transform matrix (ε) on the cost functions ($\{\|\tilde{\mathbf{y}}_i - \mathbf{r}_i\mathbf{X}\|_2^2\}_{i=1}^3$). The cost functions decrease monotonically with increasing ε; however, the decrease becomes gradual at approximately $\varepsilon = 0.2$. Assuming that the onboard sensor characteristics do not change significantly, the parameter ε is empirically set to 0.2 to allow a change of 20% from the prelaunch value. The graph on the left of Fig. 5 shows the relative spectral response between ASTER and Hyperion estimated

from observed images in the case of convex optimization. Table I shows the cost functions ($\{\|\tilde{\mathbf{y}}_i - \mathbf{r}_i\mathbf{X}\|_2^2\}_{i=1}^3$) for each multispectral band. For comparison, results obtained by the simple least-squares method are also shown on the right of Fig. 5 and in Table I. The least-squares method has the minimum cost but negative values appear in $\mathbf{R}$, which are not physically meaningful. In contrast, a physically meaningful result was obtained with the proposed method with a smaller cost function than that for the prelaunch condition. We used the result obtained by convex optimization for data fusion. Further experiments using a range of scenarios will improve the estimation accuracy of $\mathbf{R}$.

B. Data Fusion

The top rows of Table II show numerical evaluation results for the three methods. SFIM gave the best results for the first scenario and the hyperspectral-based second scenario because the spatially degraded fused image in SFIM is the original hyperspectral image as can be seen from (5). SSCN gave the best result for the multispectral-based second scenario because
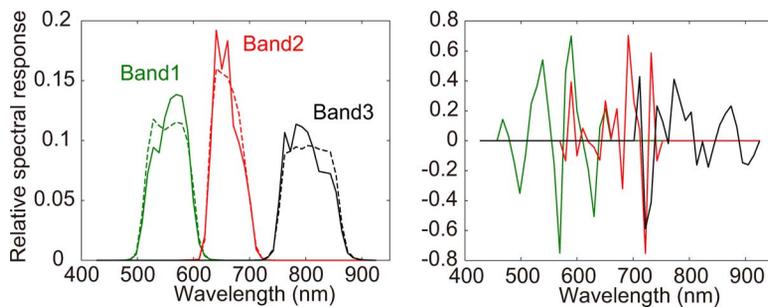
Fig. 5.   Result of estimating relative spectral response function between ASTER and Hyperion. Prelaunch (dashed line on left), convex optimization (solid line on left) and least-squares (right).

TABLE I
COST FUNCTIONS FOR EACH BAND

ASTER Band	Pre-launch	Convex Opt.	Least Squares
1	6.876	6.599	6.222
2	6.566	6.452	6.045
3	4.699	4.617	4.368

the spectrally degraded fused image in SSCN corresponds to the original multispectral image according to (7). The assumption used in SFIM and SSCN, i.e., the radiance ratio between high spatial-resolution hyperspectral and original multispectral images is the same as that between original hyperspectral and low-pass filtered multispectral images, may hold owing to the small difference between the spatial resolution of Hyperion and ASTER. CNMF gave intermediate results, which means that this method tends to inherit the characteristics of both hyperspectral and multispectral data. When the difference in the spatial resolution becomes so large that the assumption of SFIM and SSCN does not hold, CNMF can be used to enhance all hyperspectral bands with small spectral distortion owing to unmixing. Fig. 6 shows RGB color images of the original hyperspectral and CNMF fused data using bands 13 (467.5 nm), 23 (569.3 nm), and 33 (671.0 nm) for blue, green, and red, respectively. As can be seen from the enlarged figure, the image becomes clear, meaning that roads and buildings become easier to discern.

The impact of the sensor characteristics on the fused image was investigated using the prelaunch $\mathbf{R}$ and $\mathbf{S}$ instead of the corresponding matrices estimated by convex optimization. $\mathbf{R}$ and $\mathbf{S}$ are used in SSCN and SFIM, respectively. By comparing the top and bottom rows of Table II, it can be seen that the optimization of $\mathbf{R}$ improves the results of SSCN for both hyperspectral and multispectral criteria. The optimization of $\mathbf{S}$ has a positive effect on the results of SFIM in terms of multispectral-based evaluation but a negative effect in terms of hyperspectral-based evaluation. This indicates that the accuracy of the estimated spectral response function ($\mathbf{R}$) has a greater effect on the final product than that of the point spread function ($\mathbf{S}$). However, the decrease in RMSE due to the optimization of the sensor characteristic estimation is much smaller than the RMSE itself, indicating that

Fig. 6.   Original Hyperion (top) and CNMF fused (bottom) images (B: 467.5 nm, G: 569.3 nm, R: 671.0 nm).

the residual errors originate from the difference between two images caused by atmospheric effects, illumination conditions, and other sensor characteristics.

VI. CONCLUSION

Hyperspectral and multispectral data fusion was applied to real satellite hyperspectral and multispectral datasets, i.e., Hyperion and ASTER images. The preprocessing of the datasets and the estimation of the sensor characteristics were demonstrated in detail. The estimation of the relative spectral response function by convex optimization has a positive impact on the final product. The cross-calibration of sensor characteristics can be used for the data fusion of different optical images. Several fusion techniques were used for hyperspectral and multispectral data fusion in this study and their properties were investigated. Pansharpening-based methods, i.e., SFIM and SSCN, tend to be biased toward one image, while unmixing-based CNMF inherits the characteristics of both images. This paper is expected to contribute to the practical use of hyperspectral and multispectral data fusion which will enable accurate classification with

TABLE II
COMPARISONS OF AVERAGE RMSE, CC, AND SAE (IN DEGREES)

	1st scenario			2nd scenario					
				Hyperspectral			Multispectral		
	RMSE	CC	SAE	RMSE	CC	SAE	RMSE	CC	SAE
Ideal	0	1	0	0	1	0	0	1	0
SFIM	57.02	0.9972	0.7968	18.78	0.9996	0.2864	70.42	0.9965	0.5947
SSCN	80.04	0.9945	0.9758	68.43	0.9962	0.7004	0.22	1	0.0043
CNMF	69.65	0.9961	0.8094	62.59	0.9968	0.6716	33.76	0.9993	0.3486
Prelaunch									
SFIM	53.65	0.9974	0.7764	18.75	0.9997	0.2774	75.83	0.9957	0.6593
SSCN	82.23	0.9942	0.9940	71.09	0.9960	0.7004	13.93	0.9999	0.1039

fine spatial resolution and thereby enhance the applications of hyperspectral remote sensing.

The results of this study will be applicable to the Japanese next-generation earth-observing satellite, Advanced Land Observing Satellite 3 (ALOS-3), due to be launched in 2015, which will carry the Hyperspectral Imager Suite (HISUI) optical sensor, which is composed of hyperspectral and multispectral sensors [28]. HISUI will allow us to obtain hyperspectral and multispectral datasets taken over the same areas with the same observation conditions.

REFERENCES

[1] J. S. Pearlman, P. S. Barry, C. C. Segal, J. Shepanski, D. Beiso, and S. L. Carman, "Hyperion, a space-based imaging spectrometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1160–1173, Aug. 2003.

[2] Y. Yamaguchi, A. B. Kahle, H. Tsu, T. Kawakami, and M. Pniel, "Overview of advanced spaceborne thermal emission and reflection radiometer (ASTER)," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1062–1071, Jul. 1998.

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

[4] J. Vrabel, P. Doraiswamy, J. McMurtrey, and A. Stern, "Demonstration of the accuracy of improved resolution hyperspectral imagery," in *Proc. SPIE*, 2002, vol. 4725, pp. 556–567.

[5] 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.

[6] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.

[7] N. Mayumi and A. Iwasaki, "Image sharpening using hyperspectral and multispectral data," in *Proc. IEEE IGARSS*, Vancouver, Canada, 2011, pp. 519–522.

[8] G. Thoonen, Z. Mahmood, S. Peeters, and P. Scheunders, "Multisource classification of color and hyperspectral images using color attribute profiles and composite decision fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 5, no. 2, pp. 510–521, 2012.

[9] R. Dinuls, G. Erins, A. Lorencs, I. Mednieks, and J. Sinica-Sinavskis, "Tree species identification in mixed Baltic forest using LiDAR and multispectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 5, no. 2, pp. 594–603, 2012.

[10] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Dec. 2000.

[11] 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 Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.

[12] 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.

[13] T.-M. Tu, C.-L. Hsu, P.-Y. Tu, and C.-H. Lee, "An adjustable pan-sharpening approach for IKONOS/QuickBird/GeoEye-1/World-View-2 imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 5, no. 1, pp. 125–134, Feb. 2012.

[14] D. D. Lee and H. S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, pp. 788–791, Oct. 1999.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] R. E. Crippen and R. G. Blom, "Measurement of subresolution terrain displacements using SPOT panchromatic imagery," in *Proc. IEEE IGARSS*, Espoo, Finland, 1991, vol. 3, pp. 1667–1670.

[23] P. Blanc, L. Wald, and T. Ranchin, T. Ranchin and L. Wald, Eds., "Importance and effect of coregistration quality in an example of "pixel to pixel" fusion process," in *Proc. 2nd Conf. Fusion of Earth Data*, Nice, France, Jan. 28–30, 1998, pp. 67–73.

[24] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," in *Proc. IEEE ICIP*, New York, 2002, vol. 3, pp. 57–60.

[25] T. Wang, G. Yan, H. Ren, and X. Mu, "Improved methods for spectral calibration of on-orbit imaging spectrometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 3924–3931, 2010.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[27] CVX Research. [Online]. Available: http://cvxr.com/

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

Norimasa Mayumi received the M.Sc. degree in aerospace engineering from the University of Tokyo, Tokyo, Japan, in 2012. He enrolled at the University of Tokyo in 2006, where he engaged in research on space technology and remote-sensing systems.

Naoto Yokoya received the M.Sc. degree in aerospace engineering from the University of Tokyo, Tokyo, Japan, in 2010. He is working towards the Ph.D. degree in aerospace engineering at the University of Tokyo. Since April 2012, he has been a research fellow of the Japan Society for the Promotion of Science.

His research interests include hyperspectral data analysis.

Akira Iwasaki received the M.Sc. degree in aerospace engineering from the University of Tokyo, Japan, in 1987. He joined the Electrotechnical Laboratory in 1987 where he was engaged in research on space technology and remote sensing systems. He received the Doctoral degree in Engineering from the University of Tokyo in 1996.

He is currently a Professor with the University of Tokyo.