AIRBORNE UNMIXING-BASED HYPERSPECTRAL SUPER-RESOLUTION USING RGB IMAGERY

Naoto Yokoya and Akira Iwasaki

Department of Advanced Interdisciplinary Studies, The University of Tokyo, Japan

ABSTRACT

This paper presents an airborne experiment on unmixing-based hyperspectral super-resolution using RGB imagery. Preprocessing is described to ensure spatial and spectral consistency between hyperspectral and RGB images. An extended version of *coupled nonnegative matrix factorization* (CNMF) is introduced for multisensor hyperspectral super-resolution to deal with a challenging problem setting, i.e., only three spectral channels for higher spatial information and a 10-fold difference of ground sampling distance. The proposed method successfully estimated the high-spatial-resolution red-edge image. Numerical evaluation by comparing the high-spatial-resolution hyperspectral image to ground-measured spectra demonstrated recovery of pure-pixel spectra by the proposed method.

Index Terms— Hyperspectral super-resolution, data fusion, unmixing, coupled nonnegative matrix factorization (CNMF)

1. INTRODUCTION

Hyperspectral (HS) imagers generally have a larger ground sampling distance (GSD) than multispectral (MS) imagers owing to a trade-off between spatial and spectral resolutions and the signal-to-noise ratio. HS and MS data fusion enables the super-resolution of HS data [1, 2, 3, 4]. A Bayesian approach was first proposed for HS super-resolution using MS data [1, 2]. Unmixing-based HS and MS data fusion can enhance the spatial resolution of HS data with little spectral distortion [3, 4]. An unmixing-based HS and MS data fusion method, named *coupled nonnegative matrix factorization* (CNMF), is composed of alternating unmixing for two images using nonnegative matrix factorization (NMF) [5], which has recently been receiving attention for the unmixing of HS data based on a linear spectral mixture model (LSMM) to deal with severe mixtures considering nonnegativity [6, 7]. A similar approach for the super-resolution of HS data using an RGB image, which is a special case of HS and MS data fusion, was independently proposed in the field of computer vision [4].

HS and MS data fusion algorithms in remote sensing have mainly been evaluated using synthetic datasets or real datasets taken from different platforms because there has been no platform that is composed of HS and MS imagers with the trade-off between spatial and spectral resolutions. In this paper, we present an airborne experiment on unmixing-based HS super-resolution fusing with a higher-spatial-resolution RGB image taken from the same platform. To the best of our knowledge, this work is the first attempt to demonstrate the experiment of multisensor unmixing-based HS super-resolution using real remote sensing datasets taken from the same platform. We describe the preprocessing that ensures spatial and spectral consistency between two images, which is a practical issue of multisensor super-resolution. An extended version of CNMF is introduced to deal with a challenging problem setting: only three spectral channels for MS data and a 10-fold GSD difference. The reconstruction of pure-pixel spectra is numerically validated by comparison to ground-measured spectra.

2. METHODOLOGY

2.1. Flow of Preprocessing

Spatial and spectral consistency between HS and RGB images is important to generate high-spatial-resolution HS data with accurate spectra. To ensure these consistency between HS and RGB images, we adopt the following preprocessing.

1. *Smile* and *keystone*, which are distortions of spectrum images in the spectral and spatial domains, respectively, are modified using an image-matching technique [8].

2. Geometric projection is applied to two datasets using positions and orientations of the aircraft.

3. The HS image is registered to the RGB image to mitigate geometric errors. The relative spatial response function can be approximated by a Gaussian filter.

4. The HS image is converted to a reflectance image using ground-measured spectra of a reference area.

5. The estimation of relative spectral response functions (SRFs) can be formulated as a constrained least-squares problem [9, 10]. Reflectance conversion coefficients of the RGB image are simultaneously obtained.

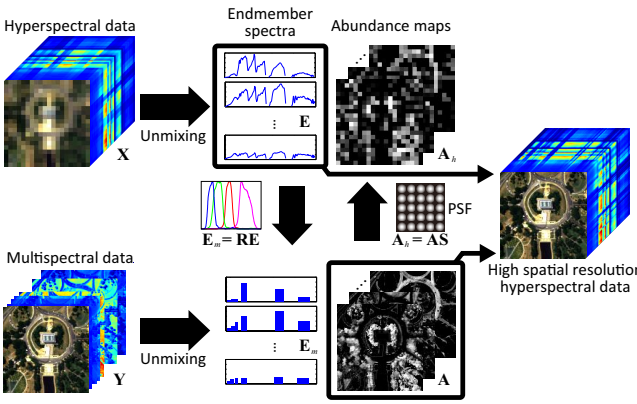


Fig. 1. Illustration of unmixing-based HS and MS data fusion.

6. The RGB image is converted to a reflectance image using the coefficients.

2.2. Unmixing-Based HS and MS Data Fusion

HS and MS data can be represented in matrix form as $\mathbf{X} \in \mathbb{R}^{L_h \times P_h}$ and $\mathbf{Y} \in \mathbb{R}^{L_m \times P_m}$, respectively. L_h and L_m denote the numbers of spectral bands, and P_h and P_m denote the numbers of pixels. $L_h > L_m$ and $P_h < P_m$ are satisfied owing to the trade-off between spectral and spatial resolutions. The high-spatial-resolution HS data is denoted as $\mathbf{Z} \in \mathbb{R}^{L_h \times P_m}$. The spectrum at each pixel is assumed to be a linear combination of several endmember spectra. Therefore, $\mathbf{Z}$ is formulated as

$$\mathbf{Z} = \mathbf{E}\mathbf{A} + \mathbf{N}, \tag{1}$$

where $\mathbf{E} \in \mathbb{R}^{L_h \times M}$ is the endmember matrix, $\mathbf{A} \in \mathbb{R}^{M \times P_m}$ is the abundance matrix, and $\mathbf{N} \in \mathbb{R}^{L_h \times P_m}$ is the residual.

Unmixing-based HS and MS data fusion yields estimates of $\mathbf{E}$ and $\mathbf{A}$ from observable $\mathbf{X}$ and $\mathbf{Y}$ to reconstruct $\mathbf{Z}$, as illustrated in Fig. 1. The low-spatial-resolution HS data and MS data can be seen as degraded versions of the high-spatial-resolution HS data in the spatial and spectral domains, respectively. Therefore, $\mathbf{X}$ and $\mathbf{Y}$ are modeled as

$$\mathbf{X} = \mathbf{Z}\mathbf{S}, \;\; \mathbf{Y} = \mathbf{R}\mathbf{Z}. \tag{2}$$

Here, $\mathbf{S} \in \mathbb{R}^{P_m \times P_h}$ is the relative spatial response transform matrix and $\mathbf{R} \in \mathbb{R}^{L_m \times L_h}$ is the spectral response transform matrix, which are obtained in the 3^{rd} and 5^{th} preprocessing steps, respectively. By substituting (1) into (2), $\mathbf{X}$ and $\mathbf{Y}$ can be approximated as two LSMMs:

$$\mathbf{X} \approx \mathbf{E}(\mathbf{A}\mathbf{S}), \;\; \mathbf{Y} \approx (\mathbf{R}\mathbf{E})\mathbf{A}. \tag{3}$$

$\mathbf{E}$ and $\mathbf{A}$ can be obtained by alternating unmixing of $\mathbf{X}$ and $\mathbf{Y}$ under the constraints of the relative sensor characteristics ($\mathbf{R}$ and $\mathbf{S}$). Here, we define $\mathbf{A}_h = \mathbf{A}\mathbf{S}$ and $\mathbf{E}_m = \mathbf{R}\mathbf{E}$ as degraded versions of $\mathbf{A}$ and $\mathbf{E}$, respectively. Spatial and spectral consistency, which is expressed as $\mathbf{R}\mathbf{X} = \mathbf{Y}\mathbf{S}$ from (3), is important in practical use and is satisfied by the preprocessing.

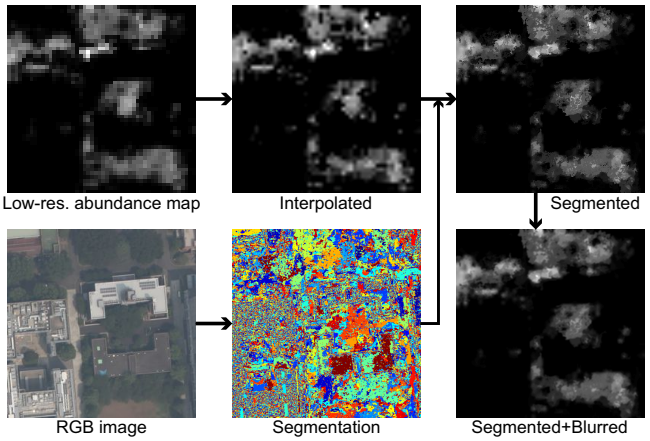


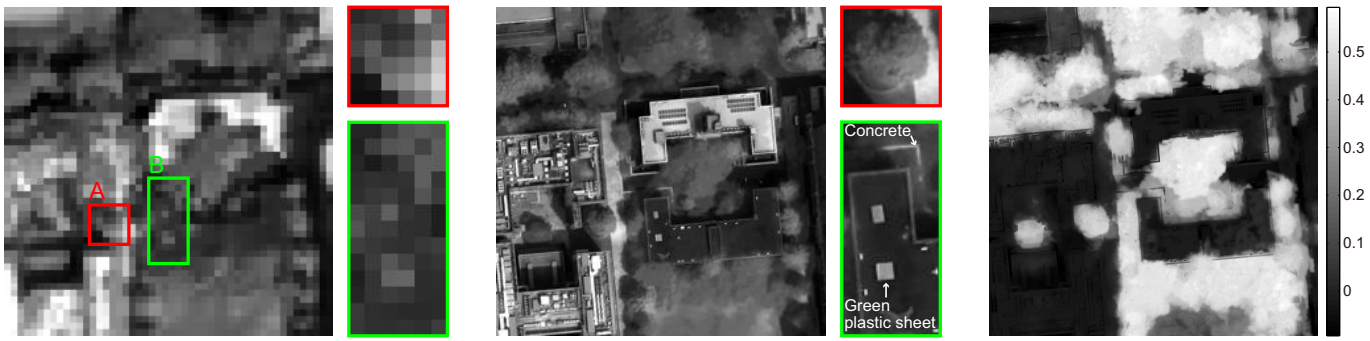
Fig. 2. Illustration of Improved initialization of high-spatial-resolution abundance map.

Since NMF has been shown to be effective for spectral unmixing satisfying physical constraints without assuming the presence of pure pixels, CNMF uses NMF to solve the two unmixing problems and couples them under the sensor characteristics. CNMF starts with NMF-based unmixing of the HS data to estimate the endmember ($\mathbf{E}$) and abundance ($\mathbf{A}_h$) matrices using its spectral advantage. NMF converges to local minima; therefore, the initialization is important. The vertex component analysis (VCA) [11] and fully constrained least-squares (FCLS) [12] methods are used for the initialization of $\mathbf{E}$ and $\mathbf{A}_h$, respectively. Next, the MS image is unmixed by NMF after initializing the endmember ($\mathbf{E}_m$) and abundance ($\mathbf{A}$) matrices using $\mathbf{E}$ and $\mathbf{A}_h$. The relative SRFs ($\mathbf{R}$) are used for initializing $\mathbf{E}_m$. The sequential unmixing for HS data is processed after initializing the abundance by $\mathbf{A}_h = \mathbf{A}\mathbf{S}$. After that, two datasets are alternately unmixed until convergence and the fused data is obtained by multiplying the endmember matrix by the high-spatial-resolution abundance matrix.

2.3. Extended CNMF

When the difference of GSDs is large and the RGB image is used for the MS data, the unmixing of the RGB image results in a severely ill-posed problem. To tackle this challenging problem, we propose an extended CNMF (ECNMF) that improves the initialization of the abundance matrix ($\mathbf{A}$).

In the ordinary CNMF, $\mathbf{A}$ is initialized as interpolated abundance maps calculated from $\mathbf{A}_h$, which may not be suitable for a large GSD difference. In this work, we introduce segmentation for better initialization of $\mathbf{A}$, as shown in Fig. 2. First, the low-spatial-resolution abundance maps ($\mathbf{A}_h$) are upscaled using image interpolation. Next, the interpolated abundance maps are converted to segmented abundance maps by integrating region-unifying segmentation of the RGB image



(a) (b) (c)

Fig. 3. 719-nm-band images of (a) HS and (b) fused data and (c) NDVI image using 719 and 691 nm images.

with the interpolated abundance maps. The average abundance value of a segment is assigned to all pixels in the segment. Finally, a Gaussian-blurred version of segmented abundances is used for initializing $\mathbf{A}$. The modified abundance fraction at the j^{th} pixel in the i^{th} band $(\hat{\mathbf{A}})_{i,j}$ is given by

$$(\hat{\mathbf{A}})_{i,j} = K \sum_{j^*=1}^{P_m} \exp\left(-\frac{g(j,j^*)^2}{2c^2}\right) \frac{1}{|\varphi_j|} \sum_{j' \in \varphi_j} (\mathbf{A})_{i,j'} \quad (4)$$

where $K = 1/\left(\sum_{j^*=1}^{P_m} \exp\left(-\frac{g(j,j^*)^2}{2c^2}\right)\right)$, φ_j is the segment where the j^{th} pixel belongs to, $g(j,j^*)$ is the Euclidean distance between the j^{th} and j^{*th} pixels, and c is a Gaussian parameter. In addition, we update only $\mathbf{A}$ in the RGB unmixing to avoid the incorrect convergence of $\mathbf{E}_m$ owing to the small number of spectral bands.

3. AIRBORNE EXPERIMENT

3.1. Image Acquisition

Airborne observation was conducted using HS and RGB cameras mounted on a small aircraft. The HS camera is a HyperSpec-VNIR-C (Headwall Photonics Inc.), which captures 128 bands in the 390–1040 nm spectral range by pushbroom imaging. The 400–800 nm spectral region was used because the remaining wavelength ranges have low accuracy of reflectance conversion. The RGB camera is an EOS 5D Mark II (Canon Inc.). A mosaic RGB image, in which subimages close to the nadir and taken by continuous shooting are merged, is used for fusion with HS data. The altitude was approximately 1000 m and the GSDs of the HS and RGB images are 2.5 and 0.25 m, respectively, after geometric projection. The dataset was acquired over the University of Tokyo campus and the neighboring urban area.

3.2. Experimental Results

We produced the 0.25-m-GSD HS data with 97 bands over the 400–800 nm spectral region by applying the ECNMF method

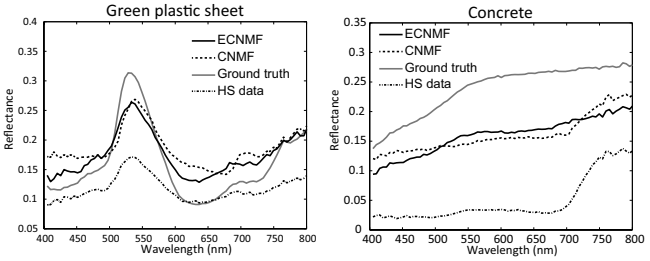


Fig. 4. Spectral comparison between fused data and low-spatial-resolution HS data as well as ground measured spectra for (left) green plastic sheet and (right) concrete shown in Fig. 3(b).

to the HS and RGB images. Figs. 3 (a)(b) show 719-nm-band images of low-spatial-resolution HS and fused data, respectively. As shown in the enlarged images, a tree crown, a green plastic sheet spread on a roof and a concrete line are identifiable with the fused data owing to the 10-fold improvement of GSD.

To evaluate the validity of the red-edge (719 nm) image of the fused data, we examine the normalized difference vegetation index (NDVI) map. Fig. 3(c) shows the NDVI map using 719 and 691 nm images. Vegetation areas show high values, whereas the green plastic sheets, which have small spectral angles compared with vegetation in the RGB image, show low values. The unobservable high-spatial-resolution red-edge image could be obtained by the unmixing-based fusion of spectral information of the HS data and spatial information of the RGB image. However, note that nonvegetation materials around vegetation are bright. It may be because the optimization of two alternating unmixings converged to local optima owing to the initial errors of the high-spatial-resolution abundance maps. Once abundances at one pixel in the RGB image are initialized to include vegetation, the NMF-based update of abundances cannot remove this effect because there is no near-infrared information that enables discrimination between vegetation and nonvegetation.

Fig. 4 shows the spectra of two objects, i.e., the green

	SAE (degree)			RMSE		
	HS	CNMF	ECNMF	HS	CNMF	ECNMF
Green plastic sheet	**8.89**	14.3	11.2	0.0759	0.0386	**0.0275**
Concrete	50.4	42.0	**27.9**	0.184	0.0769	**0.0748**

Table 1. Numerical evaluation of spectral reconstruction for three objects, i.e., green plastic sheet, compressor unit, and concrete.

plastic sheet and the concrete, for low-spatial-resolution HS and for CNMF and ECNMF fused data as well as the ground-measured spectra. The spectra of the two objects are extracted from the corresponding locations as shown in the enlarged rectangle image in Fig. 3(b). Table 1 shows the root mean squared errors (RMSEs) and spectral angle errors (SAEs) of two spectra for CNMF and ECNMF data using the ground-measured spectra as the reference data. The spectra of the low-spatial-resolution HS data appears to be degraded from the ground-measured spectra because of the mixed pixels. In contrast, the CNMF and ECNMF data recover the spectral signatures showing smaller RMSEs compared to the low-spatial-resolution HS data. ECNMF shows the best or second performance for all criteria. The deference between CNMF and ECNMF can be seen in Fig. 4. Although CNMF contains spectral artifacts caused by initialization errors of the high-resolution abundances, ECNMF successfully recover the spectral signatures more similar to the ground measurements, which proves the effectiveness of the ECNMF method.

4. CONCLUSION

We demonstrated an airborne experiment on unmixing-based HS super-resolution using an RGB image. The preprocessing is described to ensure spatial and spectral consistency between two images, which is an important condition of multisensor super-resolution. We introduced segmentation to the initialization of higher-spatial-resolution abundance maps in CNMF to deal with a severely ill-posed problem due to the 10-fold GSD difference and the lack of higher-spatial-resolution near-infrared images in the MS data. The high-spatial-resolution red-edge image could be obtained by unmixing-based HS super-resolution. The fused data were numerically evaluated using the ground-measured spectra, and we confirmed that the proposed method can recover pure-pixel spectra that are degraded because of mixed pixels in the low-spatial-resolution HS data.

5. ACKNOWLEDGEMENT

This work was supported by contract with the National Institute of Advanced Industrial Science and Technology (AIST) and KAKENHI (12J09244, 24360347).

6. REFERENCES

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

[2] 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, September 2004.

[3] 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, February 2012.

[4] R. Kawakami, J. Wright, Y.-W. Tai, Y. Matsushita, M. Ben-Ezra, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 2329–2336, 2011.

[5] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, October 1999.

[6] 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, March 2007.

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

[8] N. Yokoya, N. Miyamura, and A. Iwasaki, "Detection and correction of spectral and spatial misregistrations for hyperspectral data using phase correlation method," *Applied Optics*, vol. 49, no. 24, pp. 4568–4575, 2010.

[9] G. D. Finlayson, S. Hordley, and P. M. Hubel. "Recovering device sensitivities with quadratic programming," in *Proc. the IS and T/SID Sixth Color Imaging Conference: Color Science, Systems, and Applications*, pp. 90–95, November 1998.

[10] N. Yokoya, N. Mayumi, and A. Iwasaki, "Cross-calibration for data fusion of EO-1/Hyperion and Terra/ASTER," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 419–426, 2013.

[11] 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, 2005.

[12] 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, March 2001.