Hyperspectral and Multispectral Data Fusion:
A Comparative Review

Naoto Yokoya, *Member, IEEE*, Claas Grohnfeldt, *Student Member, IEEE*, and Jocelyn Chanussot, *Fellow, IEEE*

Abstract—In recent years, enormous efforts have been made to design image processing algorithms to enhance the spatial resolution of hyperspectral (HS) imagery. One of the most commonly addressed problems is the fusion of HS data with higher-spatial-resolution multispectral (MS) data. Various techniques have been proposed to solve this data fusion problem based on different theories including component substitution, multiresolution analysis, spectral unmixing, and Bayesian probability. This paper presents a comparative review of those HS-MS fusion techniques with extensive experiments. Ten state-of-the-art HS-MS fusion methods are compared by assessing their fusion performance both quantitatively and visually. Eight data sets featuring different geographical and sensor characteristics are used in the experiments to evaluate the generalizability and versatility of the fusion algorithms. To maximize the fairness and transparency of this comparison, publicly available source codes are used, and parameters are individually tuned for maximum performance. Additionally, the impact of spatial resolution enhancement on classification is investigated. Robustness against various factors characterizing the HS-MS fusion problem is systematically analyzed for all methods under comparison. The algorithm characteristics are summarized, and methods with high general versatility are clarified. The paper also provides possible future directions for the development of HS-MS fusion.

Index Terms—Hyperspecral and multispectral data fusion, resolution enhancement, comparative review.

I. INTRODUCTION

Upcoming spaceborne imaging spectroscopy (or hyper-spectral) missions [1]–[6] will enable the identification and discrimination of materials and the derivation of surface parameters at an accuracy level unattainable by currently operational optical broadband (multispectral) satellites. Those missions are designed and expected to enhance our understanding, possibilities, and capabilities in a wide range of applications such as monitoring and management of natural resources, the ecosystem, biodiversity, and disasters. Due to the inevitable

N. Yokoya is with the Department of Advanced Interdisciplinary Studies, the University of Tokyo, Tokyo 153-8904, Tokyo, Japan, with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Wessling, Germany, and also with Signal Processing in Earth Observation (SiPEO), Technical University of Munich (TUM), 80333 Munich, Germany (e-mail: yokoya@sal.rcast.u-tokyo.ac.jp).

C. Grohnfeldt is with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Wessling, Germany (e-mail: claas.grohnfeldt@dlr.de).

J. Chanussot is with the Grenoble Images Parole Signal Automatique Laboratory, Department of Image and Signal, Grenoble Institute of Technology, Saint-Martin-d'Hères F-38402, France, and also with the Faculty of Electrical and Computer Engineering, University of Iceland, Reykjavik 101, Iceland (e-mail: jocelyn.chanussot@gipsa-lab.fr).

trade-off between spatial resolution, spectral resolution, and signal-to-noise ratio (SNR), spaceborne imaging spectrometers are usually designed to provide data with a moderate ground sampling distance (GSD) (e.g., 30 m; see the illustration at the upper left of Fig. 1), limiting the range of potential applications.

If higher spatial resolution – possibly multispectral (MS) – data of the scene of interest is available, data fusion can be performed to generate high spatial resolution hyperspectral (HS) data (see Fig. 1). This data can be thought of as the product of a synthetic sensor featuring the high spatial resolution of the MS sensor and the high-spectral resolution of the HS sensor. It allows for various new applications potentially conducted on a global scale, which, to data, have been possible only locally with high-spatial-resolution airborne imaging systems. Such applications include high-spatial-resolution ecosystem monitoring, or the high-spatial-resolution mapping of e.g. minerals, urban surface materials, plant species, and many others. Although the number of available satellite platforms mounting both HS-MS imaging sensors is limited to date [6], the increasing number and availability of high-resolution optical satellites [7]–[9] as well as the ever improving revisit cycles allow for acquisitions of complementary HS-MS images during the same season and possibly under similar atmospheric and illumination conditions. Fig. 2 shows a scatter plot of currently operational MS satellite missions and both operating and upcoming HS satellites over GSD (y-axis) and the number of spectral bands between 0.4–2.5 μm (x-axis). Considering the currently – or soon to be – available satellite pair constellations, the potential synthesized high-resolution HS sensors would fall within the purple area in the lower right corner in Fig. 2.

A variety of HS-MS data fusion techniques have been developed in the last decade to enhance the spatial resolution of HS imagery as detailed in Section II. Because most investigators used limited data sets with slightly different evaluation methodologies, the generalizability and versatility of various HS-MS fusion methods remain unknown. Loncan *et al.* presented a comparative study of hyperspectral pan-sharpening, which is a special case of HS-MS fusion [10]. Mookambiga and Gomathi [11] and Palubinskas [12] reported reviews of resolution enhancement of HS data including HS-MS fusion; however, no comparative experiment was provided. To the best of the authors' knowledge, there is no study comparing different state-of-the-art HS-MS fusion methods with extensive experiments sufficient to draw meaningful conclusions.

In this paper, we compare ten state-of-the-art HS-MS fusion methods on a variety of data sets of different nature and
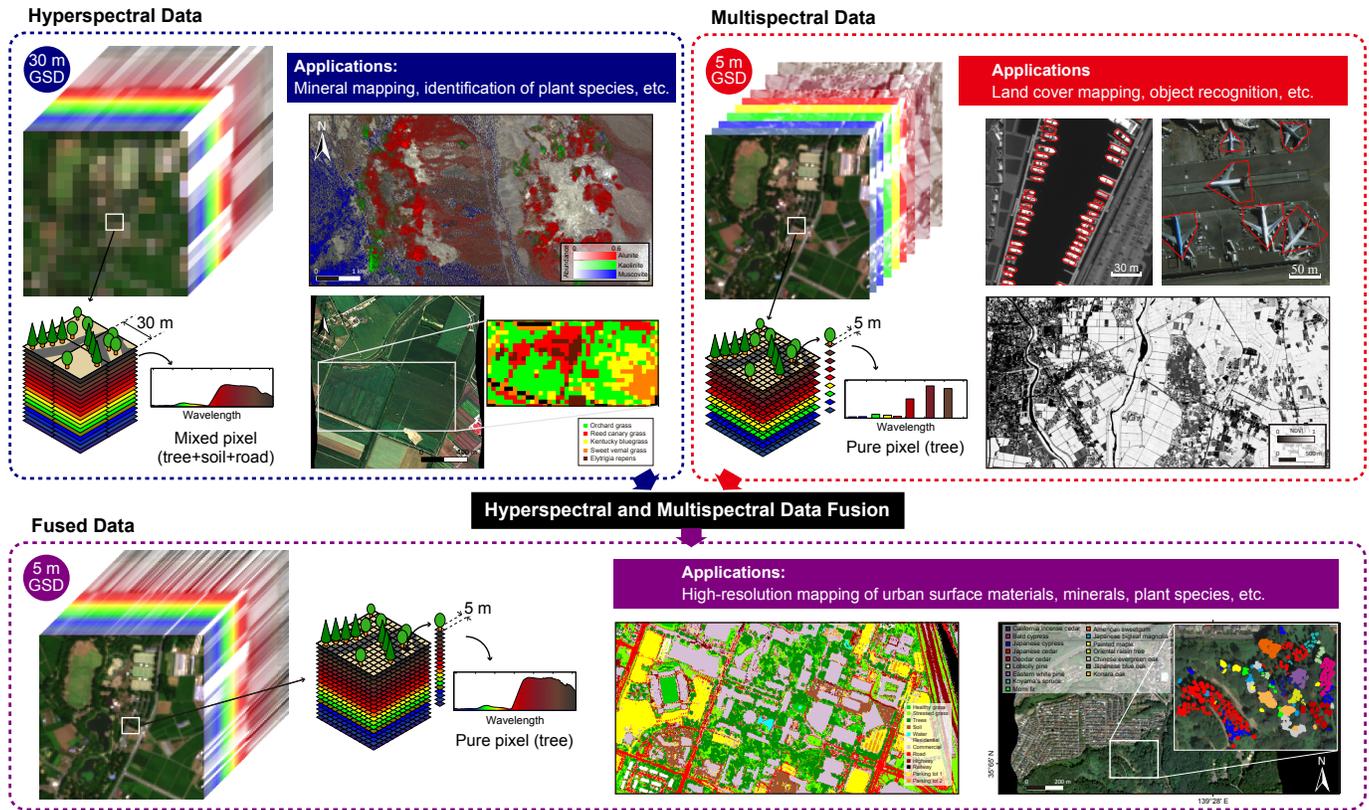
Fig. 1. Concept and benefits of HS-MS data fusion. Combining the high spatial resolution information and corresponding data analysis methods in MS data with the high spectral resolution information and corresponding material identification and discrimination capabilities opens whole new categories of applications involving high spatial high spectral resolution data possibly on a global scale.
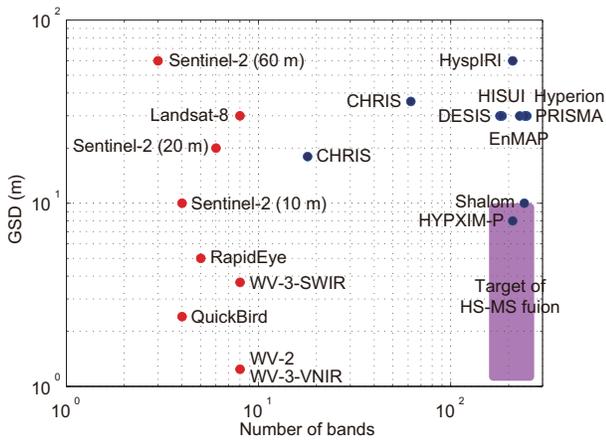

Fig. 2. GSD and the number of spectral bands in 0.4–2.5 μm of spaceborne MS (red dot) and HS (blue dot) sensors.

characteristic. The main contributions of this paper are the following:

- Objective and fair comparison: To conduct an objective and fair comparison, a total of eight data sets with diversity in scenes and fusion scenarios including but not limited to various resolution ratios and spectral overlap between the two input images were used. Publicly available source codes are used for the experiments. All methods are optimized for maximum individual performance in every experiment by extensive parameter tuning and optimal algorithm setting.

- Comprehensive evaluation methodology: Both quantitative and visual assessment of all fusion results are conducted. A complementary selection of well-established evaluation metrics ensure an objective comparison of the resolution enhancement power of all investigated algorithms and reveal individual drawbacks and advantages relative to the other methods. Furthermore, application-driven evaluation of the fused data is performed by examining the impact of all fusion results on pixel-wise classification tasks.

- Analysis of algorithm characteristics: Characteristics, strengths, and drawbacks are identified and discussed for both individual and categories of the ten state-of-the-art HS-MS fusion algorithms with different fusion scenarios. On the basis of the analyzed algorithm characteristics, we present potential future directions for the development of HS-MS fusion.

The remainder of the paper is organized as follows. Section II is devoted to an overview of the development in HS-MS fusion. Section III briefly outlines the ten state-of-the-art methods for HS-MS fusion under investigation. Section IV presents the eight data sets used in the comparative study and the evaluation methodology. Experimental results and discussions are provided in Section V. Section VI concludes the paper.

II. DEVELOPMENT IN HYPERSPECTRAL AND MULTISPECTRAL DATA FUSION

Pan-sharpening has been widely developed and used to enhance the spatial resolution of MS imagery by fusing the MS data with a corresponding higher-resolution panchromatic image [13]–[18]. A large number of pan-sharpening techniques have been developed over the past two decades; representative techniques can be roughly categorized as follows: 1) component substitution (CS) [19]–[21], 2) multiresolution analysis (MRA) [22], [23], and 3) sparse representation (SR) [18], [24], [25]. As pan-sharpening can be considered a special case of the HS-MS fusion problem, efforts have been made towards generalizing existing pan-sharpening methods for HS-MS fusion. One of the first attempts of pan-sharpening-based HS-MS fusion was presented in [26], [27] using a wavelet technique. However, the performance was highly dependent on the spectral re-sampling method, which caused difficulties in enhancing the spatial resolution of all HS band images.

Based on recent advances in pan-sharpening, more sophisticated attempts have been made to adapt pan-sharpening techniques to the general HS-MS fusion problem. Chen *et al.* proposed a framework that solves the HS-MS fusion problem by dividing the spectrum of HS data into several regions and fusing HS and MS band images in each region by conventional pan-sharpening techniques [28]. A synthetic image is generated by spectral re-sampling of the MS data and used as a high-resolution image in the spectral range that is not covered by MS bands. In [29], [30], a recent SR-based pan-sharpening method [18] was applied to HS-MS fusion, before the authors demonstrated the applicability of their *spectral grouping concept* to any pan-sharpening technique [31]. Selva *et al.* proposed a framework called *hypersharpening* that effectively adapts MRA-based pan-sharpening methods to HS-MS fusion by synthesizing a high-resolution image for each HS band as a linear combination of MS band images via linear regression [32]. It was shown that the synthesized high-resolution bands used in hypersharpening could lead to significantly better fusion results than a similar yet simpler approach in which for each HS band one high-resolution MS – namely the most correlated – band is *selected* from the available MS bands rather than synthesized.

Another popular approach fuses the HS-MS images by exploiting the inherent spectral characteristics of the scene via a subspace spanned by a set of basis vectors or spectral signatures of underlying materials (so-called *endmembers*). A Bayesian method based on a maximum a posteriori (MAP) estimation developed by Eismann *et al.*, one of the first methods published in the technical literature, was a breakthrough in terms of enhancing the spatial resolution of all HS band images using a higher-resolution data source acquired by a panchromatic or MS imager [33]–[36]. This method uses a stochastic mixing model (SMM) to estimate the underlying spectral scene characteristics and formulates a cost function that optimizes the estimated HS data relative to the input HS-MS images. It should be noted that the optimization of MAP-SMM was processed in the principal component subspace. This idea of fusing the HS-MS images based on spectral information of both input images on a subspace has been the main source of inspiration for many HS-MS fusion methods developed later [37]–[39].

The principles of spectral unmixing have been used in multi-sensor multiresolution image fusion already around 2000 [40]–[42]. Gross and Schott proposed a method that unmixes low-resolution HS data and sharpens the abundance maps by fusing it with higher-resolution panchromatic data using constrained optimization techniques [40], [41]. Zhukov *et al.* presented a technique that unmixes low-resolution images using spatial details of segmentation obtained from high-resolution images and finally sharpens low-resolution images by assigning the estimated endmember signatures to the corresponding high-resolution pixels of the segmentation map [42]. In recent years, the idea of using spectral unmixing in HS-MS fusion has been attracting considerable attention owing to its straightforward interpretation of the fusion process. Several unmixing-based methods have been proposed for HS-MS fusion, resulting in various optimization formulations which have led to state-of-the-art fusion performance [43]–[55]. The unmixing-based fusion idea aims at obtaining endmember information and high-resolution abundance matrices from the HS-MS images, respectively, under the constraints of relative sensor characteristics, such as a spectral response function (SRF) and a point spread function (PSF). The fused image can be reconstructed as the product of the two resulting matrices.

Berné *et al.* presented an HS-MS fusion method based on nonnegative matrix factorization (NMF) for mid-infrared astronomy [43]. The low-resolution HS image is unmixed by NMF, and the high-resolution abundance maps are obtained from the MS image by least squares regression; the resampled endmember signatures are fixed. Yokoya *et al.* proposed coupled nonnegative matrix factorization (CNMF) that estimates the endmember and abundance matrices via alternating spectral unmixing based on NMF under the constraints of an observation model which incorporates both the relative SRF and PSF [44], [48]. Kawakami *et al.* [46] and Wycoff *et al.* [49] used sparse regularization on the fractional abundances, assuming that the number of endmembers at each pixel is *small* compared to the number of available endmembers in the underlying endmember matrix. Akhtar *et al.* [50] proposed a method based on dictionary learning and sparse coding to obtain the endmember and abundance matrices, respectively. Lanaras *et al.* [54] introduced a projected gradient method into the alternate updates of the two underlying matrices. Veganzones *et al.* [56] demonstrated that local image processing of the unmixing-based approach has potential to mitigate the ill-posedness of the fusion problem and, therefore, improve the quality of the fusion result.

Wei *et al.* developed a Bayesian HS-MS fusion methodology using both a subspace transformation and a regularization in the fusion problem to cope with the ill-posed inverse problem. The problem formulation is based on information of the prior distribution in the observed scene such as Gaussian or sparsity promoted Gaussian [38], [57], [58]. Similar to MAP-SMM, the optimization problem was formulated on a subspace corresponding to the principal components of the input HS data. Recently, a Sylvester equation-based explicit solution was in-

tegrated into the Bayesian HS-MS fusion methodology, named *the fast fusion based on Sylvester equation* (FUSE) [59]. FUSE significantly decreased the computational complexity while achieving the same performance as the previous Bayesian HS-MS fusion algorithms. Simões *et al.* proposed a method called *HySure* based on vector-total-variation-based regularization of the spatial distribution of subspace coefficients, where the subspace can be defined either by singular value decomposition (SVD) or by endmember spectral signatures [39].

III. SELECTED HYPERSPECTRAL AND MULTISPECTRAL DATA FUSION METHODS

In this comparative study, ten HS-MS fusion methods are selected to represent the state-of-the-art in HS-MS fusion including established methods as well as recently published ones. All methods under comparison correspond to at least one of the following four categories: 1) CS, 2) MRA, 3) unmixing, and 4) Bayesian-based approaches. The unmixing and Bayesian-based methods fall into the same broader category of subspace-based methods. Fig. 3(a) depicts those categories and the correspondence with the investigated methods. Flowcharts describing the fusion processes of the CS, MRA (hypersharpening), and subspace-based approaches are illustrated in Figs. 3(b)–(d), respectively. The properties of the ten methods are summarized in Table I. *Gram-Schmidt adaptive* (GSA) [21] is adopted as the representative CS-based pan-sharpening method. Two MRA-based pan-sharpening methods, namely *smoothing filtered-based intensity modulation* (SFIM) [22] and *generalized Laplacian pyramid* (GLP) [23], are adapted to HS-MS fusion via hypersharpening. CNMF [48], HySure [39], Akhtar's [50], and Lanaras's [54] methods represent the unmixing-based approach (Akhtar's and Lanaras's methods are referred to as ECCV'14 and ICCV'15, respectively). MAP-SMM [36] and two versions of FUSE [59] are based on Bayesian probability. Appart from the Bayesian-based methods, which use principal component analysis (PCA) for subspace transformation, also the unmixing-based approaches can be considered as subspace methods, because the dimension of the space spanned by the endmembers in these methods is generally smaller than the dimension spanned by the unknown high resolution HS image. All ten methods are briefly described below.

A. GSA

Aiazzi *et al.* improved CS pan-sharpening methods by taking into account the influence of SRF on the fusion procedure [21]. In the general CS scheme, a low-resolution image is sharpened by adding spatial details obtained by multiplying the difference between a high-resolution image and a synthetic intensity component by a band-wise modulation coefficient. The improvement lies in computing the synthetic intensity component by performing a linear regression between a high-resolution image and lower-resolution bands to mitigate spectral distortion. GSA integrates this technique into the GS algorithm [20]. The most straightforward way to apply GSA to HS-MS data fusion is to construct multiple image sets for pan-sharpening subproblems where each set is composed of one

TABLE I
PROPERTIES OF HS-MS FUSION METHODS: TYPE OF SUBSPACE
TRANSFORMATION AND INDICATION OF THE DEPENDENCY ON THE
KNOWLEDGE OF SRF AND PSF (YES/NO).

Method	Subspace	SRF	PSF
GSA	GS	No	No
SFIM-HS	–	No	No
GLP-HS	–	No	Yes
CNMF	VCA	Yes	Yes
ECCV'14	Spams	Yes	No
ICCV'15	SISAL	Yes	Yes
HySure	VCA / SVD	Yes	Yes
MAP-SMM	PCA	No	No
FUSE	PCA	Yes	Yes
FUSE-S	PCA	Yes	Yes

MS band and corresponding HS bands grouped by correlation-based clustering.

B. SFIM-HS

Based on a simplified model for solar radiation and land surface reflection, SFIM sharpens the low resolution image by multiplying an upscaled lower resolution image by a ratio between a higher resolution image and its low-pass filtered version on a pixel-by-pixel basis [22]. SFIM can be performed on individual HS bands. A high resolution image can be either a selected MS band based on correlation analysis or a synthesized image obtained by a linear regression of MS bands via least squares methods, while the latter is referred to as hypersharpening in [32]. The hypersharpening technique is adopted in this work. The SFIM hypersharpening method is referred to as SFIM-HS.

C. GLP-HS

In the GLP fusion scheme [23], spatial details of each low-resolution band are obtained as the difference between a high-resolution image and its low-pass version multiplied by a gain factor, which can be computed either locally or globally. In this paper, a global gain as given in [32] is adopted. A Gaussian filter, matching the modulation transfer function (MTF) of a lower-resolution sensor, is used for low-pass filtering. As with SFIM, hypersharpening is used here to effectively adapt the GLP fusion scheme to HS-MS fusion. The GLP hypersharpening method is referred to as GLP-HS.

D. CNMF

CNMF [48] alternately unmixes the HS-MS images by NMF [60] to estimate the spectral signatures of endmembers and the high resolution abundance maps, respectively. CNMF starts by unmixing the HS image using vertex component analysis (VCA) [61] to initialize the endmember signatures. Sensor observation models that relate the two input images with the relative sensor characteristics (i.e., SRF and PSF) are built into the initialization of the MS signatures of endmembers and the low-resolution abundance maps to find better local optima. The final high-resolution HS data is obtained as the product of the spectral signatures and the high-resolution abundance maps.

(a) Categories

(b) CS-based approach

(c) MRA-based approach (hypersharpening)
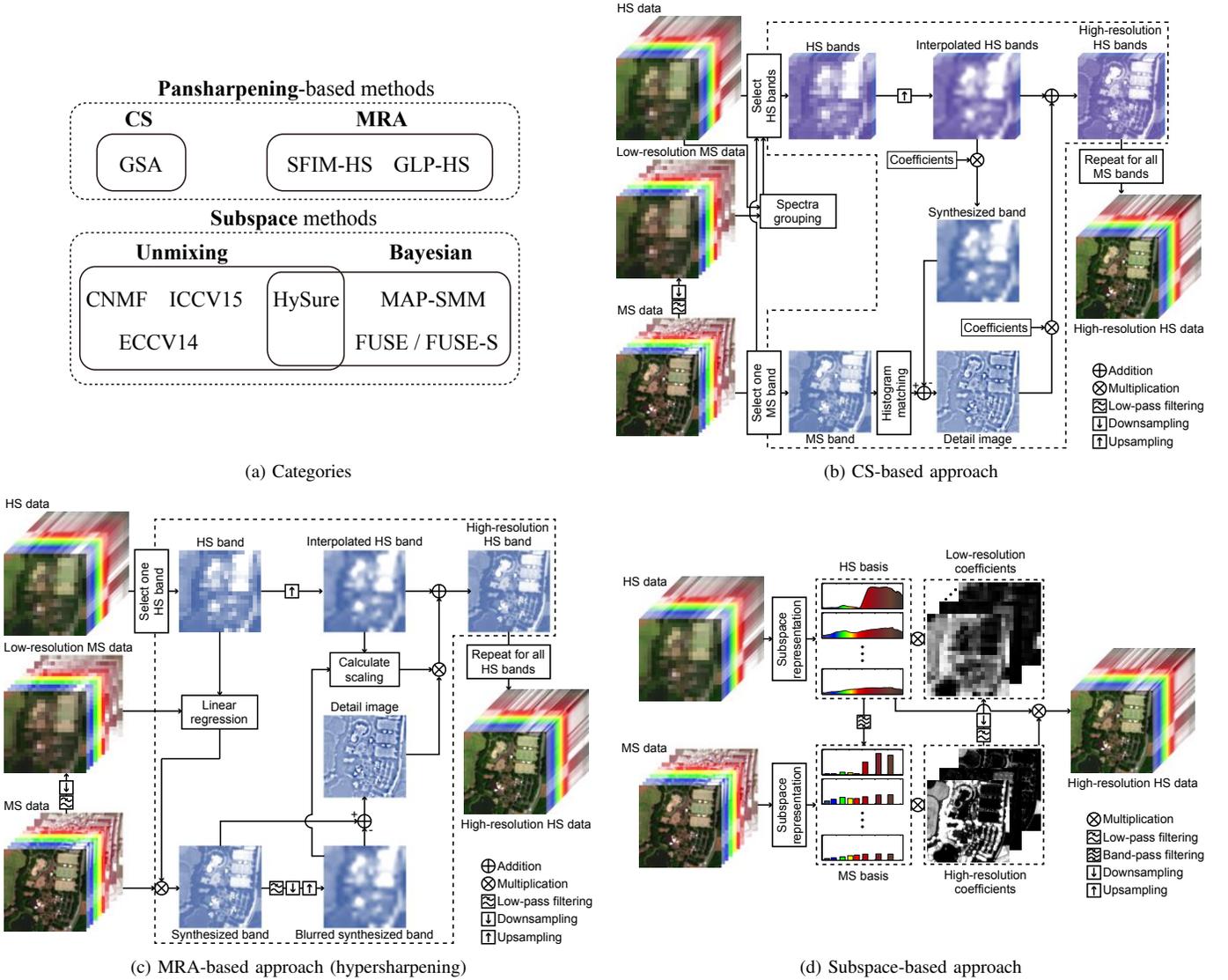
(d) Subspace-based approach

Fig. 3.   (a) Categories of HS-MS fusion methods and correspondences to the specific methods under comparison. Flowcharts of (b) CS, (c) MRA (hypersharpening), and (d) subspace-based approaches.

E. ECCV'14

Akhtar *et al*. applied dictionary learning and sparse coding to obtain the endmembers and the high-resolution abundances, respectively [50]. The spectral signatures of endmembers are obtained by the online dictionary learning method proposed by Mairal *et al*. [62]. A simultaneous greedy pursuit algorithm, named G-SOMP+, was proposed to learn a sparse code, which corresponds to the high-resolution abundances. G-SOMP+ is processed with respect to small disjoint spatial patches. The algorithm requires the availability of the relative SRF.

F. ICCV'15

Similar to CNMF, Lanaras's algorithm [54] jointly unmixes the two input images into the spectral signatures of end-members and the associated fractional abundances. A projected gradient method was proposed to alternately update the endmember signatures and the high-resolution abundances

by solving the two unmixing problems of the input HS-MS images, respectively. *Simplex identification via split augmented Lagrangian* (SISAL) [63] is used to initialize the endmembers and *sparse unmixing by variable splitting and augmented Lagrangian* (SUnSAL) [64] is adopted to obtain initial abundances. The availability of both spatial and spectral relative sensor characteristics is required to initialize the degraded versions of the endmembers and abundances.

G. HySure

HySure introduced total variation regularization into the subspace-based HS-MS fusion framework [39]. The total variation regularization is effective in preserving edges while smoothing out noise in homogeneous regions. The fusion problem is formulated as the minimization of a convex objective function with respect to subspace coefficients, which can be seen under a Bayesian approach. The subspace transformation is derived from the low-resolution HS image either by an

endmember extraction technique (e.g., VCA) or by SVD, where the former version falls into the unmixing scheme. In this paper, we adopt the unmixing version of HySure due to its empirically observed robust and high performance. The convex minimization problem is solved by the *split augmented Lagrangian shrinkage algorithm* (SALSA) [65], which is an instance of the alternating direction method of multipliers (ADMM) [66].

H. MAP-SMM

The MAP-SMM algorithm [36] adopted an SMM to estimate the underlying spectral scene statistics or, more specifically, the conditional mean vector and covariance matrix of the high-resolution HS image with respect to the MS image. The average spectrum, covariance matrix, and abundance map of each endmember are derived from the low-resolution HS image. A MAP objective function is formulated to optimize the high-resolution HS data relative to the input images based on the SMM statistics. The MAP-SMM algorithm is performed in the principal component subspace of the low-resolution HS image.

I. FUSE

FUSE utilizes a Sylvester equation to solve the maximization problem of the likelihoods obtained from the forward observation models [59]. A closed-form solution for the Sylvester equation improved computational performance. FUSE can be generalized to Bayesian estimators by integrating prior information into the fusion problem using the ADMM and the block coordinate descent method. The relative SRF and PSF are required as a priori knowledge and the optimization procedure is conducted in a principal component subspace. In this paper, naive Gaussian [59] and sparsity promoted Gaussian [38] priors are used as prior information and the corresponding algorithms are referred to as FUSE and FUSE-S, respectively.

IV. Data sets and Evaluation Methodology

This section first provides brief descriptions of the eight data sets used in the experiments. Most of the data sets are publicly available. To validate the generalizability and versatility of the HS-MS fusion algorithms, we ensure the diversity of the observed scenes (e.g., vegetation, mineral, and urban) and fusion scenarios. Next, the evaluation methodology is detailed, including simulation procedures and quality measures.

A. Data sets

Table II summarizes the main specifications of all data sets used in this work, which are described below in more detail. Moreover, RGB color composites of all data sets are shown in Fig. 4.

1) AVIRIS Indian Pines: This HS image is acquired by the airborne visible/infrared imaging spectrometer (AVIRIS) sensor over the Indian Pines test site in northwestern Indiana, USA, in 1992 [67]. The AVIRIS sensor acquires 224 spectral bands in the wavelength range from 0.4 to 2.5 μm with a full width at half maximum (FWHM) of 10 nm. The image consists of 512×614 pixels at a GSD of 20 m. We selected a 360×360 pixel size image (see Fig. 4(a)) and used 192 bands after removing bands of strong water vapor absorption and low SNRs.

2) AVIRIS Cuprite: This data set was acquired by the AVIRIS sensor over the Cuprite mining district in Nevada, USA, in 1995.[1] The entire data set comprises five reflectance images, and we used one of them saved in the file named f970619t01p02_r02_sc03.a.rfl. The full image consists of 512×614 pixels at a GSD of 20 m. We selected a 420×360 pixel size image (see Fig. 4(b)) and used 185 bands after removing noisy bands.

3) AVIRIS Moffett Field: This data set was acquired by the AVIRIS sensor over the Moffett field in California, USA, in 1997.[1] This calibrated radiance image consists of 1923×753 pixels at a GSD of 17 m. We selected a 360×360 pixel size image (see Fig. 4(c)) and used 182 bands after removing water absorption bands.

4) HYDICE Washington DC Mall: This image was taken by the hyperspectral digital imagery collection experiment (HYDICE) sensor over the Mall in Washington, DC, USA, in 1995. The HYDICE sensor acquired 210 bands between 0.4 to 2.5 μm, and the image consists of 1280×307 pixels at a GSD of 2.5 m. We selected a 420×300 pixel size image (see Fig. 4(d)) and used 191 bands after removing water absorption bands in the 0.9–1.4 μm region.

5) Hyperspec-VNIR Chikusei: The airborne HS data set was taken by Headwall's Hyperspec-VNIR-C imaging sensor over Chikusei, Ibaraki, Japan, on July 29, 2014. The data set comprises 128 bands in the spectral range from 0.363 to 1.018 μm. The scene consists of 2517 × 2335 pixels with a GSD of 2.5 m, mainly including agricultural and urban areas. We selected a 540×420 pixel size image for the experiment (see Fig. 4(e)). The data set was made available to the scientific community recently, including more detailed descriptions regarding the data acquisition and processing details, in [68].

6) ROSIS-3 University of Pavia: This data was acquired by the reflective optics spectrographic imaging system (ROSIS-3) optical airborne sensor over the University of Pavia, Italy, in 2003. It consists of 610 × 340 pixels with a GSD of 1.3 m. The 12 noisy bands have been removed, so that a total of 103 bands covering the spectral range from 0.430 to 0.838 μm are used in the experiment. The image is available as reflectance data, and a 560×320 pixel size image (see Fig. 4(f)) was selected.

7) CASI University of Houston: This data set was used in the 2013 IEEE GRSS data fusion contest [69]. The image was acquired by an ITRES CASI-1500 sensor over the University of Houston campus and its neighboring urban areas, with 144 bands covering the wavelength range from 0.364 to 1.046 μm.

---

[1][Online]. Available here: http://aviris.jpl.nasa.gov/data/free_data.html

TABLE II
DATA SET SPECIFICATIONS

Index	1 / a	2 / b	3 / c	4 / d	5 / e	6 / f	7 / g	8 / h
Site	Indian Pines	Cuprite	Moffett Field	Washington DC	Chikusei	Univ. Pavia	Univ. Houston	Rodalquilar
Original sensor	AVIRIS	AVIRIS	AVIRIS	HYDICE	Hyperspec	ROSIS-3	CASI	HyMap
Year	1992	1995	1997	1995	2014	2003	2012	2003
Data type	Radiance	Reflectance	Radiance	DN	Reflectance	Reflectance	Radiance	Reflectance
Spectral range (μm)	0.4–2.5	0.4–2.5	0.4–2.5	0.4–2.5	0.36–1.02	0.43–0.84	0.36–1.05	0.4–2.5
Bands	224	224	224	210	128	115	144	242 (126)
Used Bands	192	185	182	191	128	103	144	167
GSD	20	20	17	2.5	2.5	1.3	2.5	10 (4)
Spatial dim.	360×360	420×360	360×360	420×300	540×420	560×320	320×540	261×867
Multispectral sensor	WV-3	WV-3	QB	QB	WV-2	QB	Sentinel-2	Sentinel-2
GSD ratio	4	5	4	4	6	8	5	3
SNR (dB)	35	35	35	35	35	35	35	–

(a) AVIRIS Indian Pines

(b) AVIRIS Cuprite

(c) AVIRIS Moffett Field

(d) HYDICE Washington DC Mall

(e) Hyperspec Chikusei

(f) ROSIS-3 University of Pavia

(g) CASI University of Houston

(h) HyMap Rodalquilar

Fig. 4. Color composite images of (a) AVIRIS Indian Pines, (b) AVIRIS Cuprite, (c) AVIRIS Moffett Field, (d) HYDICE Washington DC Mall, (e) Hyperspec Chikusei, (f) ROSIS-3 University of Pavia, (g) IEEE GRSS DFC 2013 University of Houston, and (h) HyMap Rodalquilar data sets.

Fig. 5. Center wavelengths of HS bands (blue bars) and SRFs of MS imagers (red) for the (a) AVIRIS Indian Pines, (b) AVIRIS Cuprite, (c) AVIRIS Moffett Field, (d) HYDICE Washington DC Mall, (e) Hyperspec Chikusei, (f) ROSIS-3 University of Pavia, (g) CASI University of Houston, and (h) HyMap Rodalquilar data sets.

It consists of 349×1905 pixels with a GSD of 2.5 m. We selected a 320×540 pixel size image (see Fig. 4(g)).

8) HyMap Rodalquilar: The HyMap image was acquired over Rodalquilar, Spain, in June 2003 [70]. The HyMap sensor collected 126 bands in the 0.4–2.5 μm wavelength range. The scene covers a gold mining area in the Sierra del Cabo de Gata (Cabo de Gata National Park) with a GSD of 4 m. To use this data set for a realistic simulation scenario of EnMAP and Sentinel-2 data fusion, the HyMap image was spatially downgraded to a 10-m GSD and spectrally upsampled with 242 bands using linear interpolation and resampling using EnMAP SRFs. The full simulation procedure used for this test data set differs from the other seven data sets by means of the end-to-end-simulation procedure [71], [72] described in Section IV-B2. We selected a 261×867 pixel size image (see Fig. 4(h)) and used 167 bands after removing water absorption bands.

B. Simulation procedures

An accurate quality assessment of fusion data can be performed only within a simulation study [73]. The entire quality assessment procedure mainly comprises three steps: 1) simulate input HS-MS images from a reference high-resolution HS image; 2) generate a high-resolution HS image by fusing the two input images; 3) compare the fused image and the reference image.

The eight HS images described in Section III-A were used as the reference images after denoising. In the literature,

the original high-resolution HS images were often used as reference to conduct quality assessment; however, many of them include noise, which is not suitable as reference. To improve the reliability of quantitative evaluation, a denoising method in [74] is applied to the original images to increase SNRs of reference images.

For the simulation of the input images, we adopt two simulation strategies: 1) a standard simulation that considers spectral simulation, spatial simulation, and noise simulation; 2) an end-to-end simulation that takes into account the entire image acquisition and processing chain starting from raw data. The first strategy is used for the data sets #1–7 and the second strategy is used for the data set #8.

1) Standard simulation: The flow diagram of the evaluation methodology with the standard simulation is shown in Fig. 6. In the standard simulation, spectral simulation is performed to generate the MS image by degrading the reference image in the spectral domain using MS SRFs as filters. For the diversity of MS sensors, SRFs of four MS imagers were used for spectral simulation, namely, WorldView-2 (WV-2) for the data set #5, WorldView-3 (WV-3) for the data sets #1 and 2, QuickBird (QB) for the data sets #3, 4, and 6, and Sentinel-2 VNIR bands at a 10 m GSD for the data set #7 (see Table II). Fig. 5 shows overlaps of SRFs between the HS-MS imagers for all data sets. For the data sets #1, 2, 5, and 6, the MS SRFs evenly cover most of the spectral range of the HS imager. In contrast, there is no high-resolution MS image in the SWIR range for the data sets #3, 4, and 8. The latter case is more challenging and
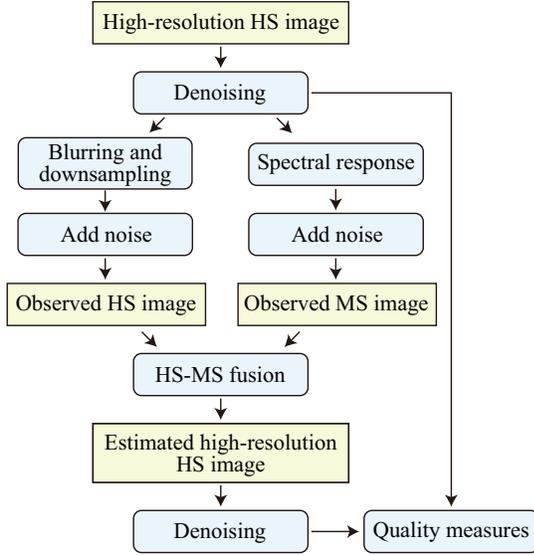
Fig. 6. Flow diagram of the evaluation methodology, including the standard simulation procedure (used e.g. in [56]).

was included in the experiment to investigate the impact of the SRF overlap between HS and MS imagers on the quality of fused data. Spatial simulation was performed to generate the low-resolution HS image using an isotropic Gaussian PSF with a FWHM of the Gaussian function equal to the resolution ratio between the GSDs of both input images. Five different GSD ratios, namely 3, 4, 5, 6, and 8, were included for spatial simulation in the standard simulation (see Table II) to simulate multiple realistic combinations of spaceborne HS-MS sensors, shown in Fig. 2. After spectral and spatial simulations, band-dependent Gaussian noise was added to the simulated HS-MS images. For realistic noise conditions, an SNR of 35 dB was simulated in all bands.

2) End-to-end simulation: On the basis of the data set #8, EnMAP and Sentinel-2 L2a (orthorectified surface reflectance data) products were simulated using the sensor end-to-end simulation tools EeteS [71] and S2eteS [72]. These tools comprise forward and backward simulators, which simulate the data acquisition procedure and the calibration and pre-processing chain, respectively, from spatially and spectrally oversampled data to the final EnMAP and Sentinel-2 products [75]. The forward simulator consists of four independent atmospheric, spatial, spectral, and radiometric modules. The spatial and spectral modules include resampling an image in the spatial and spectral domains using the sensor-specific PSFs and SRFs, respectively. The radiometric module transformed the at-sensor radiance to DN by simulating instrumental noise and calibration coefficients. The backward simulator consists of calibration modules such as non-linearity, dark current, and absolute radiometric calibration and pre-processing modules such as radiometric calibration and atmospheric correction. Compared to the standard simulation, the end-to-end simulation can generate more realistic data sets that include errors such as sensor-specific noise and residual errors of atmospheric correction. Sentinel-2 VNIR images with a GSD of 10 m (bands 2, 3, 4, and 8) were used as the MS data only,

even though the 20 m GSD SWIR images could be potentially used additionally for enhancing the EnMAP image which is of 30 m GSD.

C. Quality measures

We use the following four complementary and widely used quality measures for the quantitative fusion assessment: 1) peak signal-to-noise ratio (PSNR); 2) spectral angle mapper (SAM); 3) *erreur relative globale adimensionnelle de synthèse* (ERGAS); 4) $Q2^n$. This section briefly describes these measures and their characteristics.

Let $\mathbf{X} \in \mathbb{R}^{B \times P}$ denote the reference HS image with B spectral bands and P pixels. $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_B]^T = [\mathbf{x}_1, ..., \mathbf{x}_P]$, where $\mathbf{x}_i \in \mathbb{R}^{P \times 1}$ is the ith band ($i = 1, ..., B$) and $\mathbf{x}_j \in \mathbb{R}^{B \times 1}$ is the spectral signature of the jth pixel ($j = 1, ..., P$). $\hat{\mathbf{X}}$ denotes the estimated HS image.

1) PSNR: PSNR is used to evaluate the spatial reconstruction quality of each band. PSNR is the ratio between the maximum power of a signal and the power of residual errors. The PSNR of the ith band is defined as

$$\text{PSNR}(\mathbf{x}_i, \hat{\mathbf{x}}_i) = 10 \cdot \log_{10} \left(\frac{\max(\mathbf{x}_i)^2}{\|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_2^2 / P} \right), \quad (1)$$

where $\max(\mathbf{x}_i)$ is the maximum pixel value in the ith reference band image. A larger PSNR value indicates a higher quality of spatial reconstruction (for identical data, the PSNR is infinite). Since the PSNR is a normalized value defined at each band, it can fairly evaluate reconstruction errors of all bands without being affected by the absolute value of the data. We use the average PSNR with respect to bands for the quality index of the entire fused image. We also investigate the PSNR plot over all bands to analyze the band dependent characteristics of the fused image.

2) SAM: The SAM index [76] is commonly used to quantify the spectral information preservation at each pixel. More precisely, SAM determines the spectral similarity (or distance) by calculating the angle between two vectors of the estimated and reference spectra. The SAM index at the jth pixel is defined as

$$\text{SAM}(\mathbf{x}_j, \hat{\mathbf{x}}_j) = \arccos \left(\frac{\mathbf{x}_j^T \hat{\mathbf{x}}_j}{\|\mathbf{x}_j\|_2 \|\hat{\mathbf{x}}_j\|_2} \right). \quad (2)$$

SAM values near zero indicate high spectral quality. We use the average SAM value with respect to pixels for the quality index of the entire data set. The spatial distribution of SAM is used below to visualize location- or material-dependent the spectral quality of the fusion results.

3) ERGAS: ERGAS provides a global statistical measure of the quality of the fused data [77] with the best value at 0, which is defined as

$$\text{ERGAS}(\mathbf{X}, \hat{\mathbf{X}}) = 100d \sqrt{\frac{1}{B} \sum_{i=1}^{B} \frac{\|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_2^2}{\left(\frac{1}{P} \mathbf{1}_P^T \mathbf{x}_i\right)^2}}, \quad (3)$$

where d is the GSD ratio between the higher- and lower-spatial resolution input images, and $\mathbf{1}_P = [1, ..., 1] \in \mathbb{R}^{P \times 1}$. ERGAS calculates the band-wise normalized RMSE and multiplies it with the ratio of GSD to account for the difficulty in the fusion problem.

4) $Q2^n$: $Q2^n$ is a generalization of the universal image quality index (UIQI) [78] and an extension of the $Q4$ index [79] to hyperspectral images based on hypercomplex numbers [80]. The UIQI or the "Q" index was proposed by Wang and Bovik [78] to measure any distortion in monochromatic images as the product of three factors: loss of correlation, luminance distortion, and contrast distortion. The UIQI between one reference image band ($\mathbf{x}$) and its corresponding target image band ($\mathbf{y}$) is defined as

$$Q(\mathbf{x}, \mathbf{y}) = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2)(\bar{x}^2 + \bar{y}^2)}, \quad (4)$$

where

$$\bar{x} = \frac{1}{P}\sum_{j=1}^{P} x_j, \quad \bar{y} = \frac{1}{P}\sum_{j=1}^{P} y_j.$$

$$\sigma_x = \sqrt{\frac{1}{P}\sum_{j=1}^{P}(x_j - \bar{x})^2}, \quad \sigma_y = \sqrt{\frac{1}{P}\sum_{j=1}^{P}(y_j - \bar{y})^2},$$

$$\sigma_{xy} = \frac{1}{P}\sum_{j=1}^{P}(x_j - \bar{x})(y_j - \bar{y}).$$

UIQI can be rewritten as

$$Q(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \frac{2\bar{x}\bar{y}}{\bar{x}^2 + \bar{y}^2} \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \quad (5)$$

The three components correspond to correlation, luminance distortion, and contrast distortion, respectively. UIQI has been designed for monochromatic images. To overcome this limitation and additionally take into account *spectral* distortion, the $Q4$ index has been developed for four-band images by modeling each pixel vector as a *quaternion* [79]. $Q2^n$ further extends the $Q4$ index based on modeling each pixel spectrum ($\mathbf{x}_j$) as a hypercomplex number, namely a 2^n-ons that is represented as

$$\mathbf{x}_j = x_{j,0} + x_{j,1}i_1 + x_{j,2}i_2 + ... + x_{j,2^n-1}i_{2^n-1}. \quad (6)$$

$Q2^n$ can be computed using the hypercomplex correlation coefficient, which jointly quantifies spectral and spatial distortions [80].

V. RESULTS AND DISCUSSION

A. *Experimental setting*

To conduct a fair comparison, the fusion performance of each method was maximized by careful parameter tuning and optimal algorithm setting for each data set. The MATLAB source codes of all the algorithms are publicly available.[2] Parameter tuning and algorithm setting were performed as follows.

[2]The source codes of GSA, SFIM, GLP, and HySure are available at http://www.openremotesensing.net/index.php/codes/11-pan-sharpening. The CNMF code is available at http://naotoyokoya.com/Download.html. The ECCV'14 code is available at http://www.csse.uwa.edu.au/~ajmal/code/HSISuperRes.zip. The ICCV'15 code is available at https://www1.ethz.ch/igp/photogrammetry/people/lanaras. The MAP-SMM code is available in [36]. The FUSE code is available at http://wei.perso.enseeiht.fr/publications.html

- The key process of hypersharpening, as used for the methods SFIM-HS and GLP-HS, is to synthesize a high-resolution image for each HS band using a linear regression of MS bands. We consider two options for a linear regression, 1) least squares and 2) nonnegative least squares, and adopt an option that shows better results for each data set. As a result, the least squares method was used for the data sets #5, 6, and 7, and the nonnegative least squares method was used for the data sets #1, 2, 3, 4, and 8. This suggests that the nonnegative least squares method is suitable for data sets with limited SRF overlaps, while the least squares method performs better for data sets with good SRF overlaps (see Fig. 5). Note that band selection that assigns the most correlated MS band to each HS band was also tested along with the synthesis of a high-resolution band and compared with hypersharpening; the performance of hypershaprpening was better for all tested data sets, which confirmed the effectiveness of hypersharpening for pan-sharpening based methods.

- We used estimated SRFs obtained by the nonnegative least squares method [81] for all the methods that require SRF as input (see Table I). Additional constraints, such as smoothness of SRFs or consistency with pre-launch SRFs, may be useful for estimating the relative SRF from real data sets [39], [82]; however, the nonnegative least squares method was adopted because it empirically led to better fusion results.

- A Gaussian filter with an FWHM at GSD was used for all the methods that require PSF except HySure. HySure uses a PSF estimation method described in [39].

- The number of endmembers (k) is a key parameter for unmixing-based methods. In this paper, k was set to 30 for CNMF, ICCV'15, and HySure as this number demonstrate stable and consistently competitive performance in all experiments and data sets.

- For ECCV'14, the number of atoms in the dictionary (k) was set to the number of HS bands (B_h), the number of atoms selected in each iteration of G-SOMP+ (L) was set to 20, and a parameter for modeling error (η) in dictionary learning was set to 10^{-3} after parameter tuning. Note that dictionary learning doesn't work well if we set η as a very small value (e.g., 10^{-8} as is in the original code).

- ICCV'15 originally used SISAL to initialize endmember spectra. We also investigated its performance with the use of VCA and adopted the one that showed better results for each data set. As a result, SISAL was used for the data sets #2, 5, 6 ,7, and 8, while VCA was used for the data sets #1, 3, and 4.

- There are several key parameters for MAP-SMM, such as the number of subspace (k), the number of endmembers (n_e), and the number of mixture classes (n_m). In this paper, we set $k = 6$, $k_e = 4$, and $k_m = 35$, which are experimentally confirmed as proper numbers. Note that if the number of endmembers is large, the number of sample spectra for a certain endmember decreases. This results in a singular endmember covariance matrix, which makes the SMM process impossible.

TABLE III
ALGORITHM-SPECIFIC MAIN PARAMETERS.

Method	Parameter
GSA	−
SFIM-HS	−
GLP-HS	−
CNMF	$k = 30$
ECCV'14	$k = B_h,\ L = 20,\ \eta = 10^{-3}$
ICCV'15	$k = 30$
HySure	$k = 30$
MAP-SMM	$k = 6,\ n_e = 4,\ n_m = 35$
FUSE / FUSE-S	$6 \le k \le 15$

- The subspace dimension (k) is a critical parameter for FUSE / FUSE-S. We select the optimal number for each data set maximizing the fusion performance. Depending on the data set, parameter tuning has shown peak performance usually within a range of $6 \le k \le 15$.

B. Visual analysis

First, we visually evaluate the quality of the fused images. In Figs. 7–14, color composite images of the reference and input HS images with three selected bands are shown in the first row and those of the fusion results are presented in the first column of the second to eleventh rows. Most of the fusion methods considered in this paper provide good results and it is hard to visually discern the differences between the reference and fused images. Therefore, the errors of the fused images are visualized with three options, namely error (difference) images of color composites after contrast stretching, pixel-wise root-mean-square error (RMSE) images, and SAM images from the second to fourth columns, respectively, in the second to eleven rows of Figs. 7–14. The color-composite error images visualize the spatial distribution of errors in three selected bands, where gray pixels mean no fusion error and colored pixels indicate local spectral distortion. It is important to reveal error characteristics of individual bands because they have a major impact on many applications based on spectral indices and band ratios that select very few specific bands. The RMSE images visualize the magnitude of the error at each pixel spectrum. The SAM images visualize the spatial distribution of spectral angle errors. From the results, we make six observations.

1) The methods operating on subspaces (GSA, CNMF, ECCV'14, ICCV'15, MAP-SMM, FUSE, and FUSE-S) show material dependent errors resulting in visible spatial structures in the color-composite error images. This is because the accuracy of the fusion process can be determined at each pixel by the reconstruction error of each spectral signature via subspace representation.
2) The color-composite error images of GSA, SFIM-HS, GLP-HS, and MAP-SMM include mainly random patterns as to be seen in Figs. 7 and 8. Since those four methods do not include implicit denoising in their algorithms, noise can be contained in the fused images if the denoising post-processing fails to remove all the noise. The denoising post-processing is important for these four methods. Its impact is investigated below in Section IV-D.

3) SFIM-HS, MAP-SMM, and FUSE show block error patterns in Figs. 11, 12, and 13. This can be explained by the fact that those three methods directly use information obtained from the low-resolution HS image or adopt local processing in a shifting window. SFIM-HS estimates a high-resolution image by modulating an nearest-neighbor-upscaled low-resolution image. MAP-SMM adopts statistical information of spectral signatures based on classification of the low-resolution HS image in the fusion process. FUSE optimizes the high-resolution coefficient matrix on a block-by-block basis [59].
4) GSA, SFIM-HS, and GLP-HS show edge error patterns, particularly in Figs. 9 and 10. This is because those three methods sharpen the low-resolution image by adding spatial details obtained either by calculating the difference between the high-resolution image and the upscaled low-resolution synthetic image (in the case of GSA) or by multiresolution analysis of the high-resolution image (in the cases of SFIM-HS and GLP-HS).
5) ECCV'14 shows comparatively poor results even after extensive parameter tuning. This is particularly evident in the tests with data sets #3, 4, and 8 (see Figs. 9, 10, and 14), where the SRF overlaps between the two imagers are limited. This lack in performance can be tracked down to the estimation of high-resolution abundances which is only based on the MS image and sparsity prior information. The estimation result is likely unreliable because this sparse regression problem is severely ill-posed.
6) In SAM images, all methods show spatial patterns and they are correlated to each other for some of the data sets (e.g., the data sets #5–8). SAM values are generally high for pixels that correspond to materials with low intensity (DN, radiance, or reflectance) values due to low SNRs.

C. Quantitative evaluation

We further investigate the quality of the entire resolution-enhanced HS data via quantitative measures. The quality of the fused images was assessed by using the four quality measures described in Section IV-C. Tables IV–VI show the quality measures of the fused data sets obtained by the ten methods for the data sets #1–3, #4–6, and #7–8, respectively. The best results are shown in bold type and the second best results are indicated with underlines. From the tables, we make five observations.

1) The hypersharpening methods generally demonstrate competitive performance. In particular, when the MS SRFs cover the spectral range of the HS imager in a balanced manner and the number of MS bands is small (e.g., the data sets #5, 6, and 7), the hypersharpening methods outperform many of the other methods.
2) Several subspace-based methods (CNMF, ICCV'15, HySure, and MAP-SMM) show good performance for the data sets #1 and 2, outperforming the hypersharpening methods. This is because the estimation accuracy
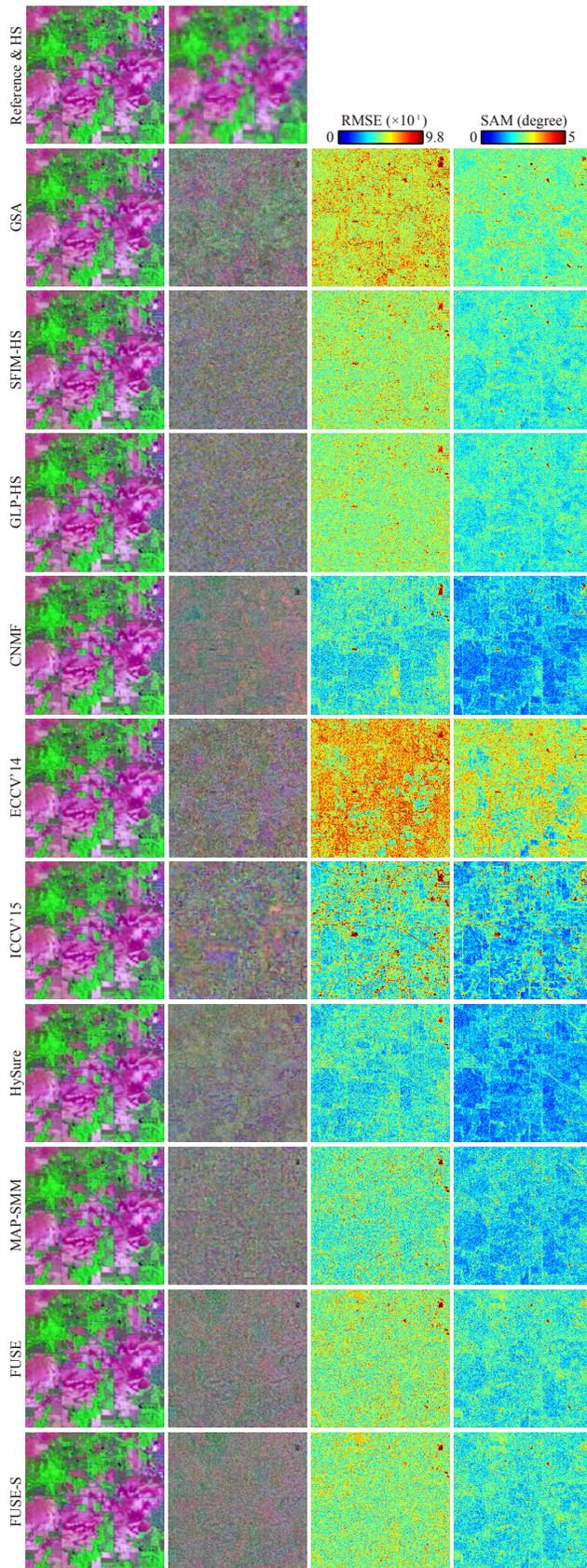
Fig. 7. AVIRIS Indian Pines data (data set #1): (1st row) Color composite images of reference and input HS images are displayed for a 240×240 pixels sub-area using 2.20, 0.80, and 0.46 μm for red, green, and blue, respectively. (2nd-11th rows) Color composites of fused images (1st column), error images relative to the reference data visualized by differences of color composites (2nd column), RMSE (3rd column), and SAM images (4th column).
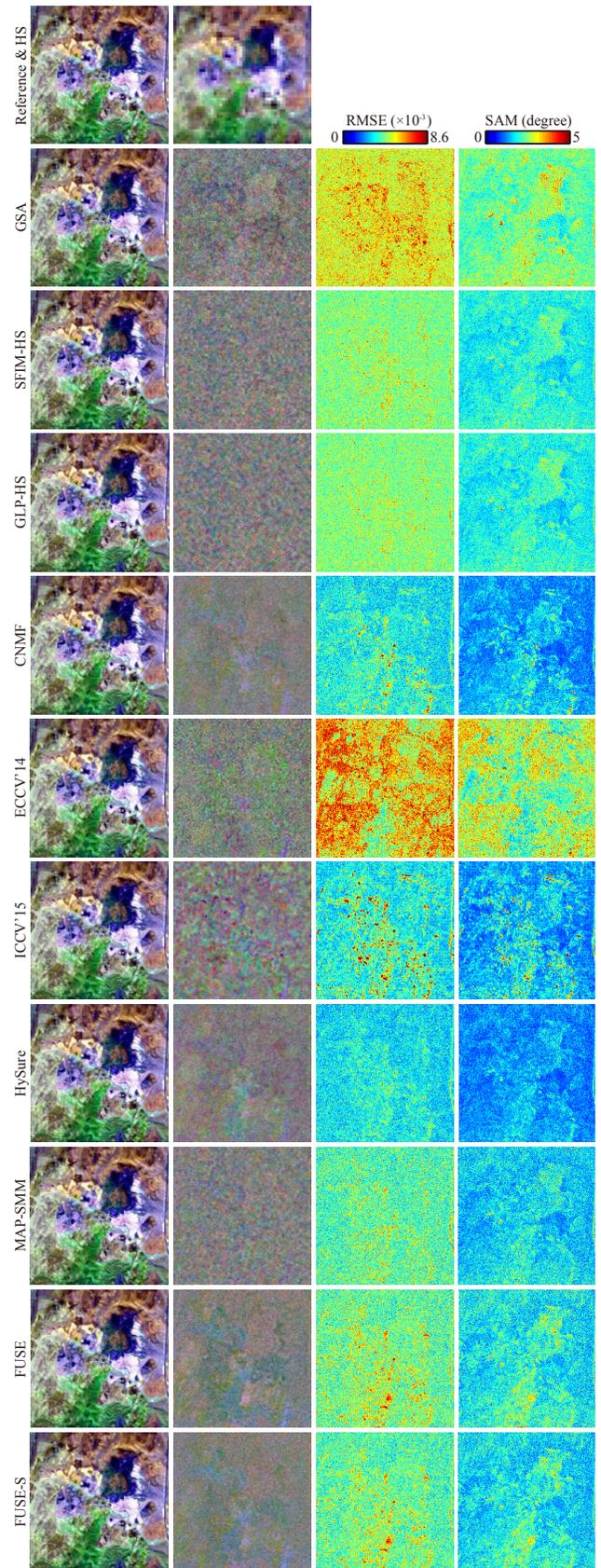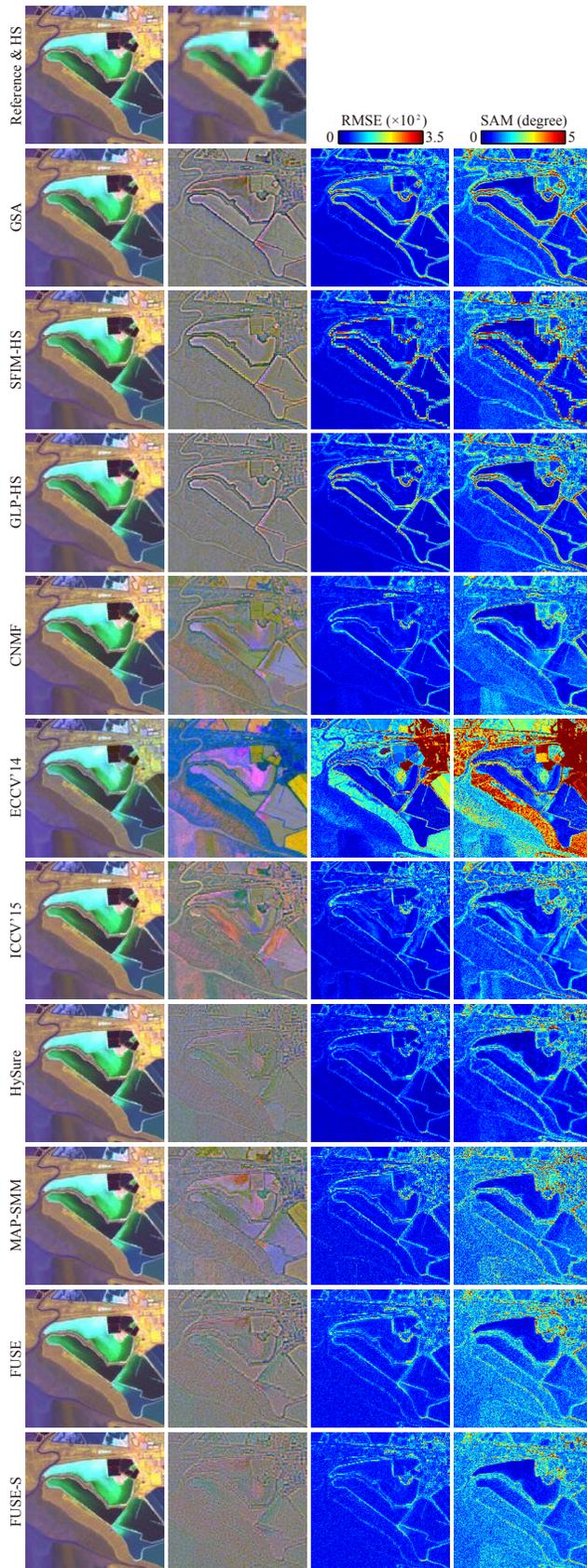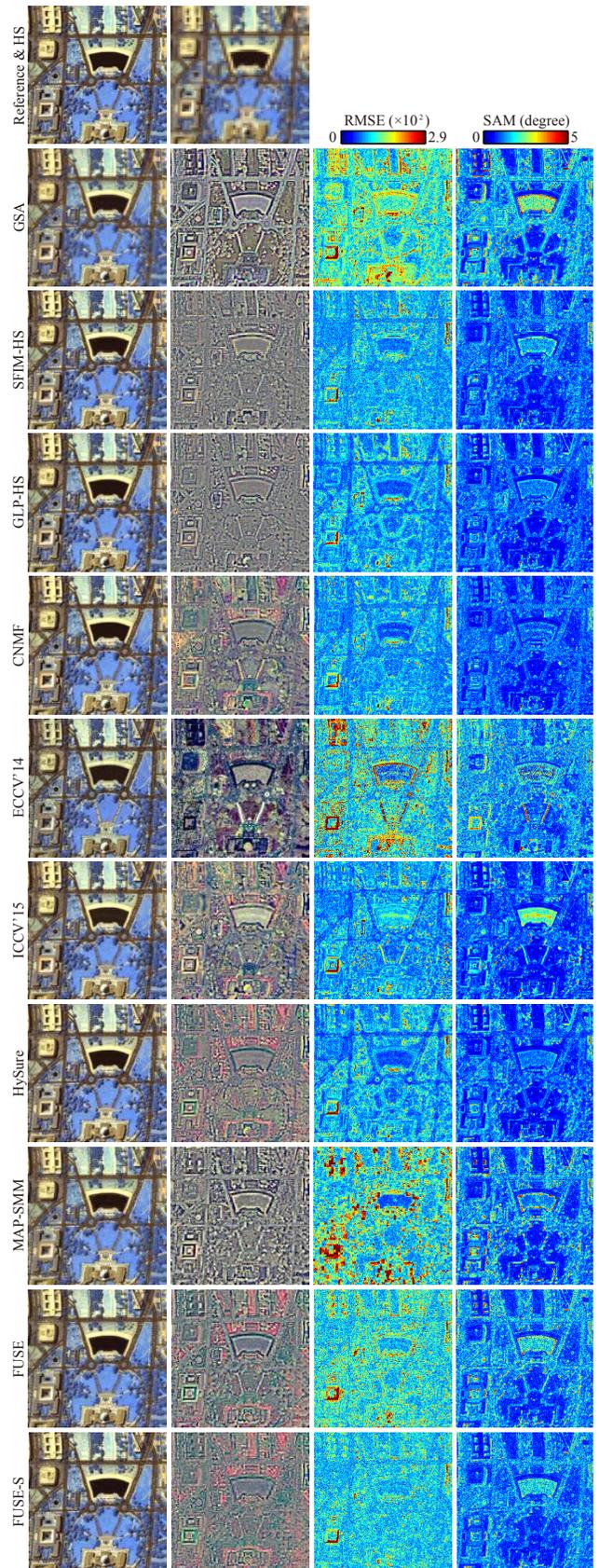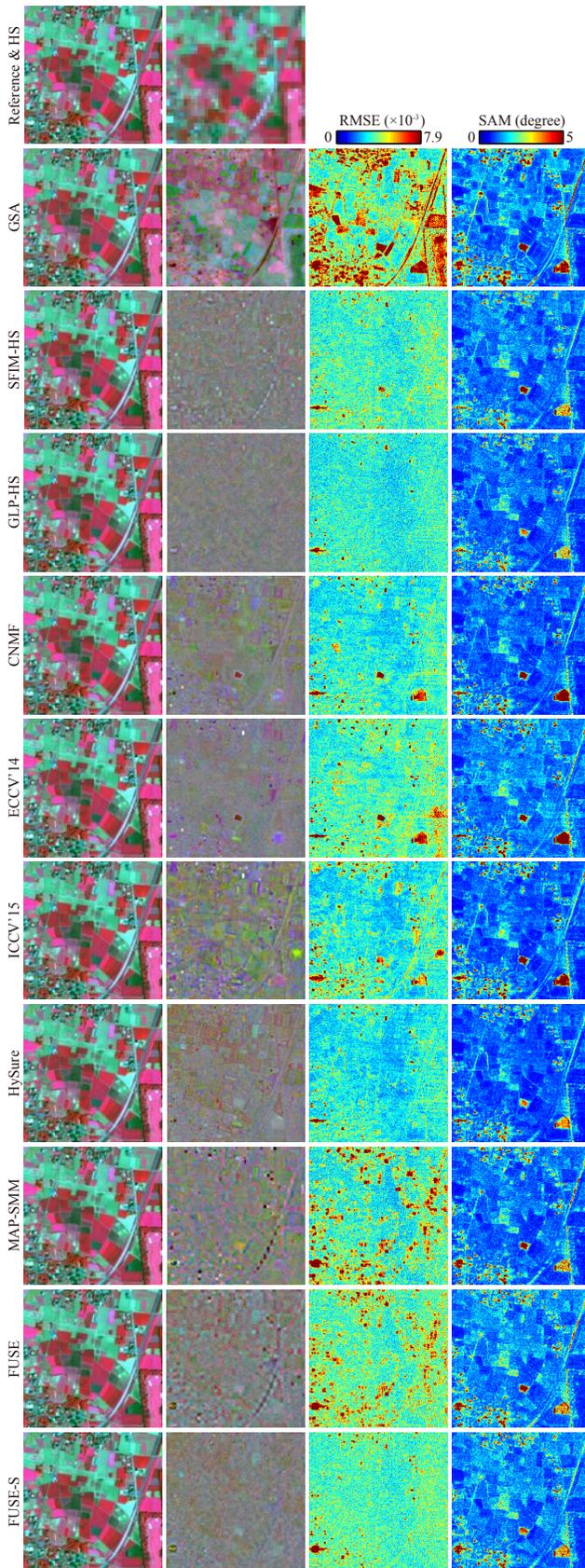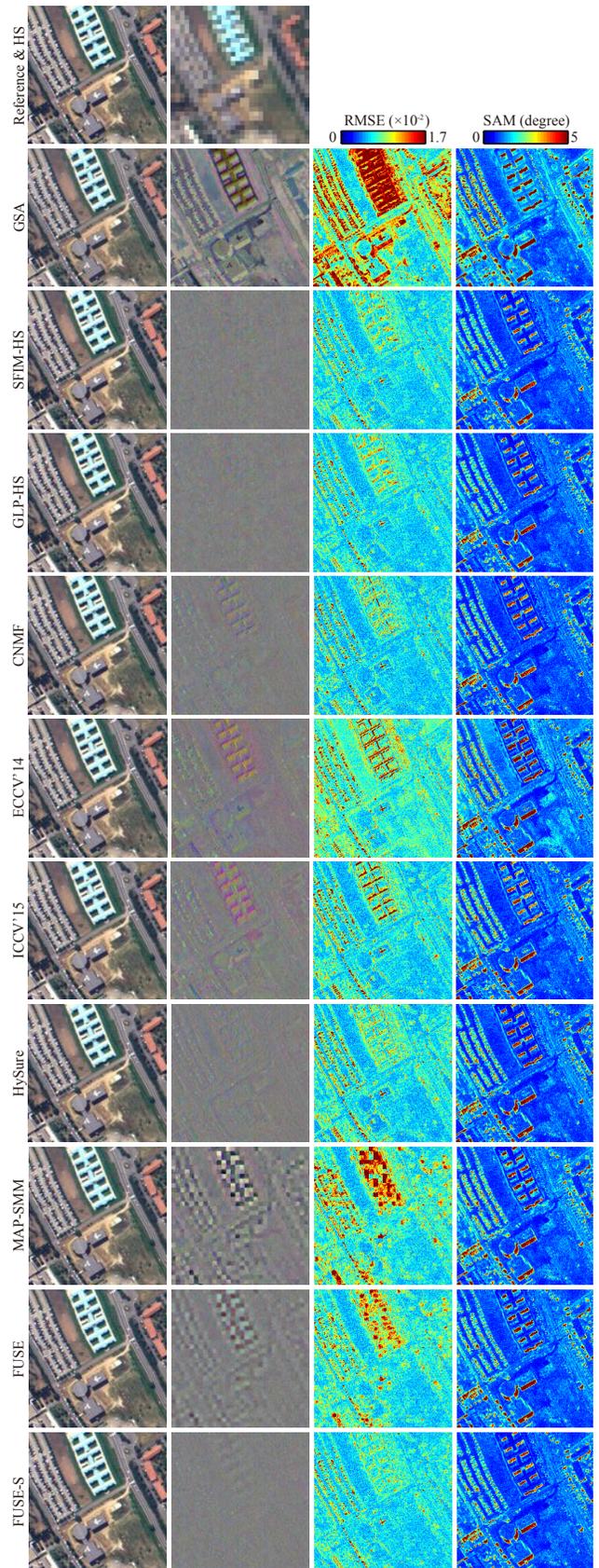
Fig. 8. AVIRIS Cuprite data (data set #2): (1st row) Color composite images of reference and input HS images are displayed for a 240×240 pixels sub-area using 2.20, 1.6, and 0.57 μm for red, green, and blue, respectively. (2nd-11th rows) Color composites of fused images (1st column), error images relative to the reference data visualized by differences of color composites (2nd column), RMSE (3rd column), and SAM images (4th column).
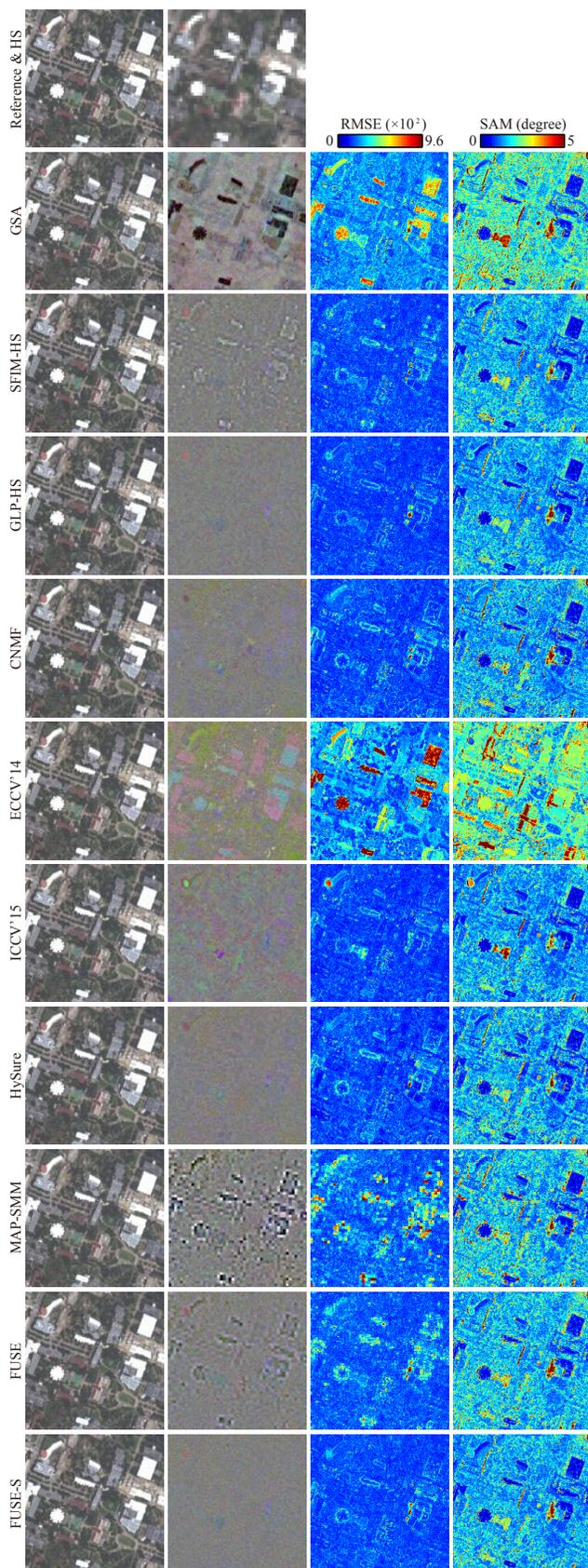
Fig. 9. AVIRIS Moffett Field data (data set #3): (1st row) Color composite images of reference and input HS images are displayed for a 240×240 pixels sub-area using 1.24, 0.86, and 0.40 μm for red, green, and blue, respectively. (2nd-11th rows) Color composites of fused images (1st column), error images relative to the reference data visualized by differences of color composites (2nd column), RMSE (3rd column), and SAM images (4th column).
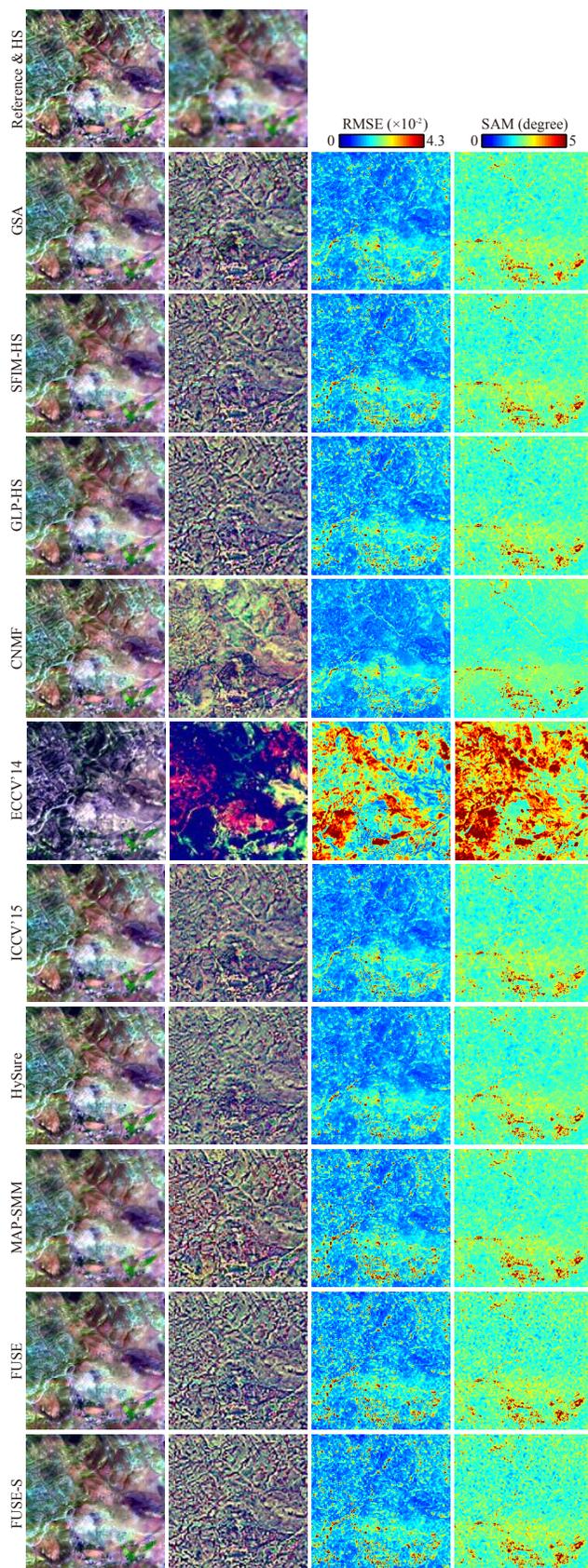
Fig. 10. HYDICE Washington DC Mall data (data set #4): (1st row) Color composite images of reference and input HS images are displayed for a 240×240 pixels sub-area using 2.20, 1.60, and 1.01 μm for red, green, and blue, respectively. (2nd-11th rows) Color composites of fused images (1st column), error images relative to the reference data visualized by differences of color composites (2nd column), RMSE (3rd column), and SAM images (4th column).

Fig. 11. Hyperspec Chikusei data (data set #5): (1st row) Color composite images of reference and input HS images are displayed for a 240×240 pixels sub-area using 0.80, 0.67, and 0.55 μm for red, green, and blue, respectively. (2nd-11th rows) Color composites of fused images (1st column), error images relative to the reference data visualized by differences of color composites (2nd column), RMSE (3rd column), and SAM images (4th column).
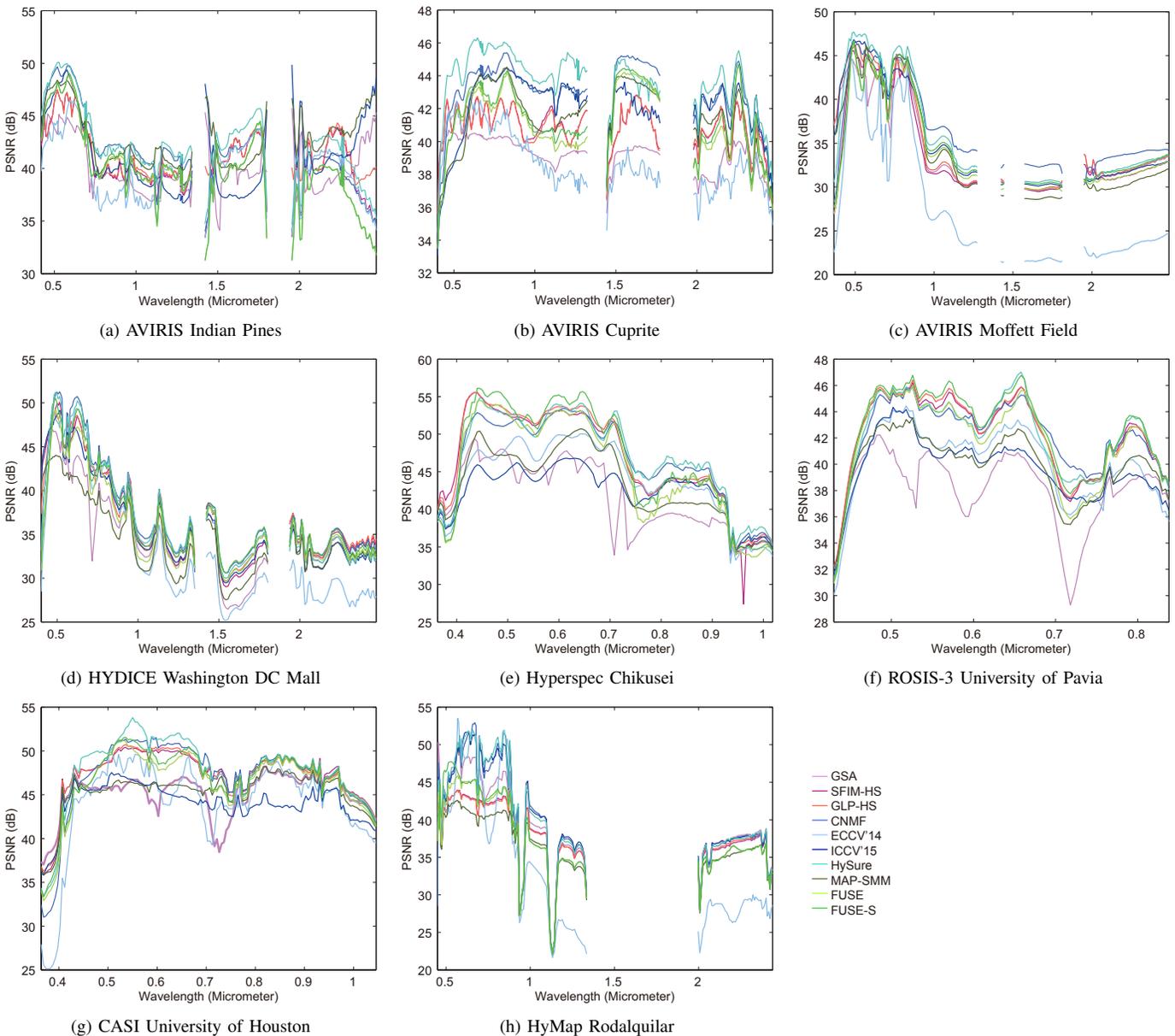
Fig. 12. ROSIS-3 University of Pavia data (data set #6): (1st row) Color composite images of reference and input HS images are displayed for a 240×240 pixels sub-area using 0.67, 0.57, and 0.46 μm for red, green, and blue, respectively. (2nd-11th rows) Color composites of fused images (1st column), error images relative to the reference data visualized by differences of color composites (2nd column), RMSE (3rd column), and SAM images (4th column).

Fig. 13. CASI University of Houston data (data set #7): (1st row) Color composite images of reference and input HS images are displayed for a 240×240 pixels sub-area using 0.65, 0.55, and 0.45 μm for red, green, and blue, respectively. (2nd-11th rows) Color composites of fused images (1st column), error images relative to the reference data visualized by differences of color composites (2nd column), RMSE (3rd column), and SAM images (4th column).

Fig. 14. HyMap Rodalquilar data (data set #8): (1st row) Color composite images of reference and input HS images are displayed for a 240×240 pixels sub-area using 2.20, 1.20, and 0.66 μm for red, green, and blue, respectively. (2nd-11th rows) Color composites of fused images (1st column), error images relative to the reference data visualized by differences of color composites (2nd column), RMSE (3rd column), and SAM images (4th column).

Fig. 15. Comparison of PSNRs between the ten HS-MS fusion algorithms for the (a) AVIRIS Indian Pines, (b) AVIRIS Cuprite, (c) AVIRIS Moffett Field, (d) HYDICE Washington DC Mall, (e) Hyperspec Chikusei, (f) ROSIS-3 University of Pavia, (g) CASI University of Houston, and (h) HyMap Rodalquilar data sets.

of high-resolution subspace coefficients or abundances increases as the number of MS bands increases.

3) The performance of the unmixing-based methods (except ECCV'14) is good and stable as a whole. In particular, the unmixing-based methods outperform the Bayesian methods when the SRF overlap is low (e.g., the data sets #3, 4, and 8). This result suggests that it is better to process the data fusion in the original spectral space rather than the projected space for the spectral ranges without MS bands.

4) The unmixing-based methods (except ECCV'14) show their advantage in the spectral quality measured by SAM for all data sets.

5) The performance of ECCV'14 is limited when there is a

large spectral range covered by the HS sensor but not by the MS bands (data sets #3, 4, and 8), which confirms the fifth observation in the visual analysis.

Fig. 15 shows PSNR plots over the HS bands. PSNR plots reveal reconstruction errors in each band separately. Many of the considered methods show comparable results. As an overall trend, the PSNR values are very good (> 40 dB) in the spectral range covered by the MS bands, whereas they dramatically decrease in the spectral ranges not covered by the high-resolution sensor (e.g., the SWIR range in the data sets #3, 4, and 8). GSA shows low PSNR values in spectral regions missing MS bands (see Fig. 15(e)(f)(g)). This is because those spectral bands are usually lowly correlated with the corresponding high-resolution images. For instance,

GSA is simply applied to HS-MS fusion by solving multiple pan-sharpening problems after grouping HS bands with respect to one MS band at a time based on correlation analysis. There is room for improvement by performing more detailed spectral grouping (e.g., [28], [31]) and synthesizing high-resolution images as hypersharpening. The method of grouping remains an open issue. ECCV'14 shows good PSNR values in the spectral ranges covered by MS bands. On the other hand, they dramatically decrease in the other spectral ranges, which is possibly caused by inaccurate estimation of high-resolution abundances.

D. Robustness analysis

The robustness of all methods under comparison against noise, SRF overlap, GSD ratios, and misregistration are systematically analyzed as follows.

1) Noise robustness: Fig. 16 shows the impact of denoising to the overall fusion quality using PSNR and SAM as quality measures. It demonstrates that the denoising post-processing consistently improves the performance of GSA, SFIM-HS, GLP-HS, and MAP-SMM while neither improving nor worsening the fusion results produced by the other methods. This observation can be explained by the fact that CNMF, ECCV'14, ICCV'15, HySure, FUSE, and FUSE-S all operate on a subspace. This meas that the dimension of their fusion result is only as large as the dimension of the corresponding underlying subspace. In case of unmixing based methods, this dimension is given by the pre-defined number of endmembers. Any denoising method reduces the dimension of the supposedly high dimensional noisy data to some extend. Now, subspace-based methods naturally imply denoising powers already. Therefore, depending on the denoising technique at hand, post-denoising does not affect those fusion results. On the other hand, this observation affirms that denoising post-processing is essential in the evaluation methodology to fairly compare resolution-enhancement performance of different approaches, whenever HS-MS fusion simulations take into account additive noise for input images.

2) SRF overlap: To further investigate the impact of SRF overlap on the quality of fused images, HS-MS fusion was performed on the AVIRIS Cuprite data using the following three scenarios of WV-3 MS bands: 1) VNIR-SWIR 16 bands 2) VNIR 8 bands; and 3) SWIR 8 bands.

Figs. 17(a) and 17(b) presents PSNR and SAM, respectively, with the three aforementioned scenarios of MS bands. As can be expected, in the second and third scenarios, the fusion performances of all methods decrease compared to those in the first scenario. In particular, the Bayesian methods and ECCV'14 show significant degradation in both PSNR and SAM, whereas the pan-sharpening-based methods, CNMF, ICCV'15, and HySure present relatively stable results. This finding suggests that HS-MS fusion should be performed on the original hyperspectral features rather than linearly transformed features such as principal components, if the SRF overlap is relatively little.

3) GSD ratio: We examine the robustness of the ten HS-MS fusion algorithms against the GSD ratio using the Hyperspec Chikusei data set. Figs. 17(c) and 17(d) show PSNR and
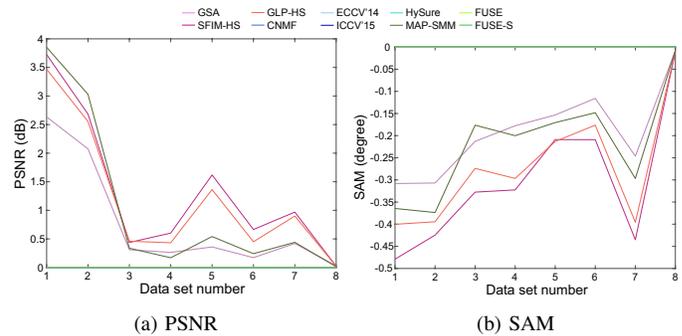

(a) PSNR (b) SAM

Fig. 16. Impact of denoising on quality assessment of fused data using (a) PSNR and (b) SAM.

SAM for the ten fused images with three cases of the GSD ratio, i.e., three, six, and ten. As the GSD ratio increases, the fusion performances decrease in general. For both quality measures, the hypersharpening methods, CNMF, and HySure show the most robust results.

4) Misregistration: The robustness against misregistration is analyzed using the Hyperspec Chikusei data set. We simulated two cases of global misregistration between the input HS-MS images by shifting one of them three and six pixels in the higher resolution in both horizontal and vertical directions. The fusion results are compared to those without misregistration. We assumed that the reference image is co-registered to the MS image.

Figs. 17(e) and 17(f) show PSNR and SAM for all methods with the perfect image registration and the two cases of misregistration. The fusion performances significantly decrease as the degree of misregistration increases, which suggests the importance of image registration as preprocessing in practice. Noticeably, ECCV'14 and ICCV'15 show relatively robust performances against misregistration. Those two algorithms optimize the high-resolution abundances based on the MS image at the end of the fusion procedure. Therefore, the fused images are spatially more consistent with the MS image. On the other hand, the other fused images are biased to the HS image, leading to higher sensitivity to misregistration given that the reference image is co-registered to the MS image. The reasons for that are two-fold: the other subspace-based algorithms optimize the fused images so that the spatially degraded versions of the fused data are consistent with the input HS image; the pan-sharpening-based algorithms add spatial details to the input HS image.

E. Impact of HS-MS fusion on classification

Finally, the quality of the fused images is indirectly validated via pixel-wise classification, which is one of the most relevant topics in the analysis of HS remote sensing data. Since the University of Pavia and University of Houston data sets have been widely used for validating classification performance in the HS image processing community owing to the availability of ground truth information, we perform classification on those two data sets. The classification performances obtained using the fused images are compared to

(a) Impact of SRF overlap on PSNR

(c) Impact of GSD ratio on PSNR

(e) Impact of misregistration on PSNR

(b) Impact of SRF overlap on SAM

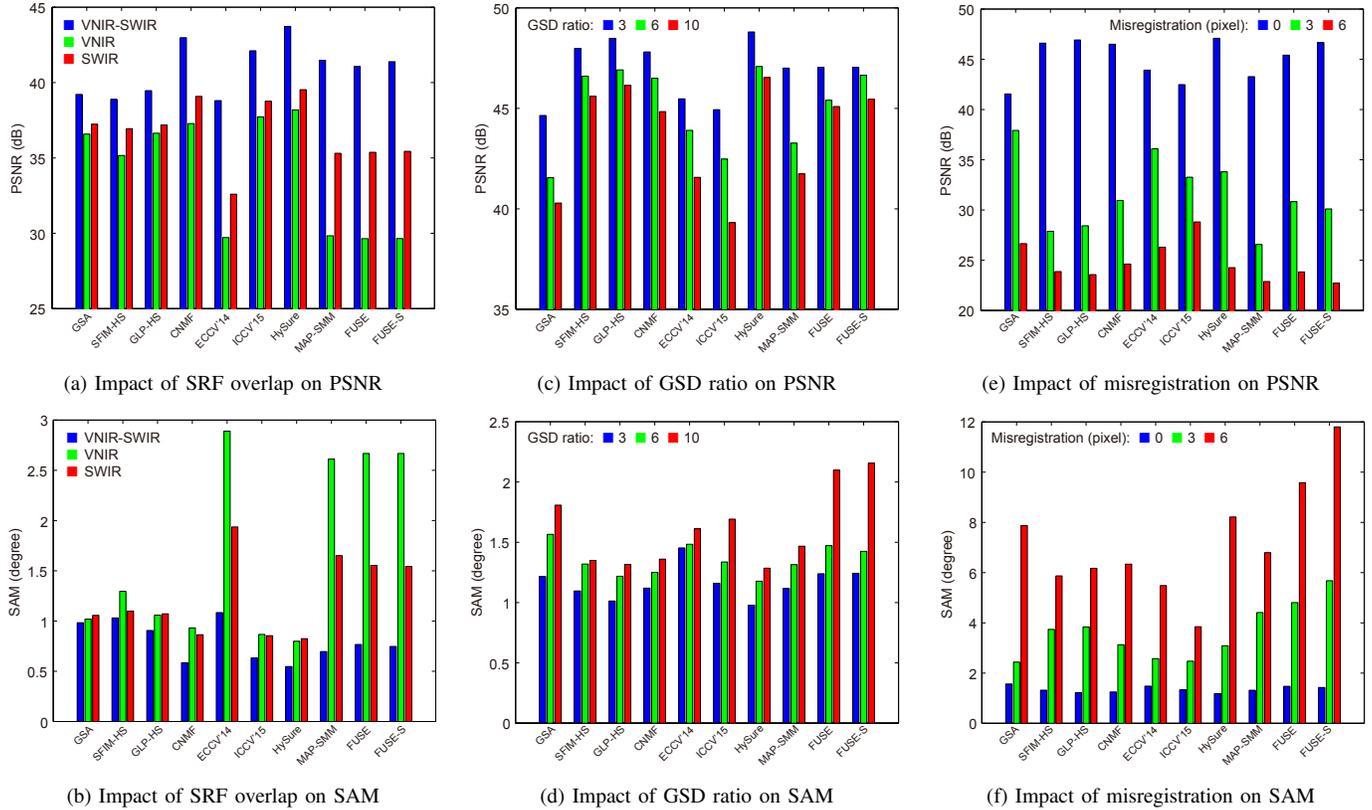(d) Impact of GSD ratio on SAM

(f) Impact of misregistration on SAM

Fig. 17. Robustness of all methods under comparison against SRF overlap (left: (a) and (b)), GSD ratio (middle: (c) and (d)), and misregistration (right: (e) and (f)) measured by PSNR (top: (a), (c), and (e)) and SAM (bottom: (b), (d), and (f)) for Hyperspec Chikusei data. Fusion results with three scenarios indicated by blue, green, and red bars are compared for each sensitivity analysis.

those obtained using the reference and input HS-MS images. Support vector machines (SVMs) [83] and rotation forests (RoFs) [84] were adopted as classifiers because of their good performance for classifying HS data with a small number of training samples [85], [86]. The class names and numbers of training and test samples are summarized in Table VII. The classification performance is quantitatively validated using the overall accuracy (OA), the average accuracy (AA), and the kappa coefficient (κ). Since training samples were randomly selected, we repeated the classification test ten times.

Table VIII shows the results of average OA, AA, and κ under different trials for the University of Pavia and University of Houston data sets. Fig. 18 presents the classification maps of one trial for the two data sets obtained by RoF using the reference, HS, MS, and HySure-fused images together with the test sample maps. From the results, we make four observations:

1) The classification accuracies obtained by RoF using most of the fused images (except ECCV'14) are higher than those using the input HS-MS images, which demonstrate the benefit of HS-MS fusion in the presented classification tasks. This is also visually demonstrated in Fig. 18. The classification maps of HySure show much better spatial details compared with those of the input HS images. Although the classification maps of the input MS images are spatially detailed, the classification maps of HySure demonstrate better material discrimination

capabilities than those of the MS images as shown in, for example, meadows and self-blocking bricks in the University of Pavia data set.

2) The average OA, AA, and κ are largely consistent with the fusion quality assessment results reported in Tables V and VI. For example, HySure led to very good classification results whereas ECCV'14 resulted in poor classification accuracies for both data sets.

3) Noticeably, for some methods the relative classification performance does not confirm the ranking in the statistical quantitative assessment tests above. GSA, for instance, showed a high potential for classification purposes even though it was not competitive in the previously conducted quantitative assessment tests.

4) Some of the fused images led to higher classification accuracies than those of the reference images. This is mainly due to the denoising effect of HS-MS fusion or denoising post-processing. Higher SNRs of the fused images can mitigate salt-and-pepper misclassified pixels.

It should be noted that classification-based validation of HS-MS fusion results does not reflect the absolute quality of fused data; firstly, because ground-truth information is only available for a comparatively small spatial subset of the data, and, secondly, because of the many factors and steps in the data preprocessing chain that may influence the classification capabilities relative to the ground-truth data. Characteristics

of applications and data analysis methods highly influence the validation results. It is possible, for instance, to design a fusion method specifically for one classification, application or data analysis scenario for which it incorporates as much information in its underlying mathematical fusion model as available. Dispite its presumably outstanding performance in that scenario, it would certainly fail in a more diverse and comprehensive testing setup like the one presented in this paper. If one is interested in finding a fusion method whose high-resolution data output is particularly suitable for SVM or RoF-based classification, the performance comparison presented in this Section V-E provides a good starting point for a more-indepth study tailored to the scenario of interest.

VI. CONCLUSIONS AND FUTURE LINES

In this paper, we presented a comparative study of HS-MS fusion methods for the resolution enhancement of HS imagery. Ten state-of-the-art HS-MS fusion methods based on four different approaches (CS, MRA, unmixing, and Bayesian) were analyzed, assessed and compared using visual, quantitative, and classification-based means of evaluation. Publicly available program codes were used while performing careful parameter tuning at optimal individual algorithm settings. To evaluate the generalizability and versatility of HS-MS-fusion performance, eight simulated data sets were used in the experiments based on various representative fusion scenarios. The overall fusion characteristics of all methods under comparison regarding performance and robustness are summarized in Table IX. Those results are primarily derived from a quantitative assessment of the fusion results under various fusion scenarios, including variations in

- imaging sensor types and sensor combinations,
- spatial parameters such as the resolution ratio between the two input images and the fusion performance robustness against misregistration in the data, and
- spectral parameters such as the spectral coverage of the HS sensor (i.e., VNIR or VNIR-SWIR) and the relative spectral overlap with the MS sensor (i.e., different numbers and spectral responsivities of MS bands).

It can be noted that, in the majority of fusion scenarios, the ranking of the ten analyzed methods, which results from the individual performance evaluations, does not change significantly. In particular, the visual and quantitative statistical comparisons are mostly consistent. Remarkably, the classification-based quality assessment revealed that some of the fusion methods, which showed good stable and robust overall performances throughout all tests based on statistical metrics, did not achieve competitive scores in the classification-based comparisons. The method which showed the most consistent and high performance in all tests including visual, statistical and classification-based assessment is the *Hyperspectral Superresolution* (HySure) method by Simões *et al.* [39]. In the overall comparison, HySure is closely followed by CNMF and, further, by the hypersharpaning adapted pan-sharpening-based methods GLP-HS and SFIM-HS.

More specific findings regarding the individual methods under comparison can be summarized as follows.

- Algorithms that were originally designed for pan-sharpening and adopted to the HS-MS fusion problem by hypersharpening are well competitive with methods that are specifically developed for HS-MS data fusion. In particular, the methods SFIM-HS and GLP-HS showed stable results above average in the majority of fusion problems, especially when the spectral range of the HS sensor is widely covered by MS bands.
- Unmixing-based methods (CNMF and HySure) demonstrated good and stable overall fusion performance throughout all tests. Particularly, outstanding results are obtained when the SRF overlap is limited, e.g., if a VNIR-SWIR HS sensor is combined with a purely VNIR MS sensor. For both methods, caution is advised in the high-resolution abundances estimation step; outstanding results could be achieved due to the minimization of the unmixing reconstruction errors with respect to both HS-MS images rather than only the MS image (as done e.g. by ECCV'14). This difference made CNMF and HySure preserve the consistency between the HS and fused images better than ECCV'14. A clear advantage of HySure over CNMF was observed in the classification tests, in which HySure maintained its previously shown performance superiority while CNMF revealed a flaw not visible in the tests based on established assessment criteria.
- Bayesian methods based on linear subspace transformations (FUSE and FUSE-S) showed good quantitative and visual performance especially if the SRF overlap between the two sensors is not too limited. A degradation in performance can be observed when the spectral range of the HS sensor is largely uncovered by MS bands. The classification capabilities are similar to, yet a little worse than, those attained by CNMF.
- Additive noise in the simulated input HS-MS images has a significant impact on the fusion quality as some methods perform denoising implicitly while others do not. Therefore, denoising is necessary for the evaluation methodology to compare the resolution-enhancement performance of different algorithms fairly. This is rarely done in the literature, which is one of the reasons why many comparisons are not doubtlessly representative.
- Most of the presented methods (except ECCV'14) demonstrated the benefit of HS-MS fusion in classification applications. That is, the classification accuracy of the fusion results exceed the accuracy achieve by the individual input HS-MS images.

It is worth noting that hypersharpening and unmixing-based methods presented similar performances in terms of quantitative quality measures, while different characteristics of reconstruction errors were observable in the visual analysis. A possible future direction for further performance improvement lies in developing hybrid approaches that combine the advantages of different classes of methods [87], such as MRA-based hypersharpening and unmixing-based approaches [88]. Comprehensive modeling and compensation of realistic noise and registration errors are still widely ignored in the fusion

community [89]. One reason for the superior performance of HySure is the algorithm's implicit denoising capacity. However, there is still much potential for robustness improvement when accounting for more error sources, such as partial coverage of one input image by clouds or seasonal differences between the acquisitions. In addition, an optimal design of HS-MS sensors capable of acquiring high-resolution HS data can be of interest for future optical Earth observation missions, as such data opens new fields of applications on a global scale, which have been explorable only locally by airborne sensors so far. Further research with real data sets will be needed to verify the practicability of HS-MS fusion methods for the upcoming satellite missions.

TABLE VIII
OVERALL CLASSIFICATION ACCURACY FOR THE ROSIS-3 UNIVERSITY OF PAVIA AND CASI UNIVERSITY OF HOUSTON DATA SETS.

Data set	University of Pavia						University of Houston					
Classifier	SVM			RoF			SVM			RoF		
Accuracy	OA	AA	κ	OA	AA	κ	OA	AA	κ	OA	AA	κ
Reference	69.51	76.92	63.59	70.35	76.94	64.77	74.54	72.95	71.51	78.47	77.06	75.86
HS	70.03	70.45	64.22	65.85	67.15	59.27	73.65	72.55	70.56	67.56	66.11	63.84
MS	64.66	71.82	57.92	64.79	72.39	58.30	71.71	70.59	68.40	71.70	70.12	68.33
GSA	**75.35**	**80.77**	**70.63**	80.41	84.44	76.61	72.20	70.81	68.90	77.07	76.18	74.34
SFIM-HS	71.91	77.82	66.49	74.79	80.18	70.02	71.96	70.35	68.65	73.29	71.67	70.10
GLP-HS	74.34	80.03	69.39	75.67	81.22	71.10	72.21	70.38	68.89	74.33	72.79	71.26
CNMF	69.32	76.65	63.62	74.27	78.95	69.35	72.27	71.15	69.01	74.57	73.10	71.54
ECCV'14	63.70	72.80	57.19	64.62	71.69	58.17	70.13	69.05	66.64	69.64	67.58	66.01
ICCV'15	72.05	78.24	66.72	76.14	80.85	71.59	75.10	**73.60**	72.15	74.92	73.17	71.93
HySure	74.02	80.04	69.10	**80.45**	**84.94**	**76.62**	75.18	73.32	72.22	79.48	77.91	77.03
MAP-SMM	71.95	79.04	66.68	74.31	80.15	69.43	71.09	70.12	67.71	74.14	72.88	71.05
FUSE	71.10	77.76	65.54	74.17	79.44	69.27	71.67	70.66	68.36	73.47	72.14	70.32
FUSE-S	69.42	76.94	63.64	73.72	79.04	68.77	71.29	69.94	67.90	73.16	71.73	69.96

Fig. 18. Classification maps of reference, HS, MS, and HySure images obtained by rotation forests for (top) University of Pavia and (bottom) University of Houston data sets.

TABLE IX
PROS AND CONS OF ALL METHODS UNDER COMPARISON. ONE, TWO, OR THREE ● SIGNS MEAN LOW, MEDIUM AND HIGH.

Method	Performance			Robustness				
	General	Spatial	Spectral	Noise	SRF	GSD	Misreg.	Material
GSA	●●	●●	●●	●	●●●	●●	●	●
SFIM-HS	●●	●●●	●●●	●	●●●	●●●	●	●●●
GLP-HS	●●●	●●●	●●●	●	●●●	●●●	●	●●●
CNMF	●●●	●●●	●●●	●●	●●●	●●●	●	●●
ECCV'14	●	●●	●	●●	●●	●●	●●	●
ICCV'15	●●	●●	●●●	●●	●●●	●●	●●	●●
HySure	●●●	●●●	●●●	●●●	●●●	●●●	●	●●
MAP-SMM	●●	●●	●●	●	●●	●●●	●	●●
FUSE	●●	●●	●●	●●	●●	●●	●	●●
FUSE-S	●●	●●●	●●	●●	●●	●●	●	●●

References

[1] L. Guanter, H. Kaufmann, K. Segl, S. Förster, C. Rogaß, S. Chabrillat, T. Küster, A. Hollstein, G. Rossner, C. Chlebek, and et al., "The EnMAP spaceborne imaging spectroscopy mission for earth observation," *Remote Sens.*, vol. 7, pp. 8830–8857, 2015.

[2] A. Iwasaki, N. Ohgi, J. Tanii, T. Kawashima, and H. Inada, "Hyperspectral imager suite (HISUI) - Japanese hyper-multi spectral radiometer," in *Proc. IEEE IGARSS*, Jul. 2011, pp. 1025–1028.

[3] P. Stefano, P. Angelo, P. Simone, R. Filomena, S. Federico, S. Tiziana, A. U. C. Vincenzo, N. Acito, D. Marco, and et al., "The PRISMA hyperspectral mission: Science activities and opportunities for agriculture and land monitoring," in *Proc. IEEE IGARSS*, Jul. 2013, pp. 4558–4561.

[4] R. Green, G. Asner, S. Ungar, and R. Knox, "NASA mission to measure global plant physiology and functional types," in *Proc. Aerospace Conference*, Mar. 2008, pp. 1–7.

[5] S. Michel, P. P. Gamet, and M. J. Lefevre-Fonollosa, "HYPXIMA hyperspectral satellite defined for science, security and defense users," in *Proc. IEEE WHISPERS*, Jun. 2011, pp. 1–4.

[6] A. Eckardt, J. Horack, F. Lehmann, D. Krutz, J. Drescher, M. Whorton, and M. Soutullo, "DESIS (DLR Earth sensing imaging spectrometer for the ISS-MUSES platform," in *Proc. IEEE IGARSS*, Jul. 2015, pp. 1457–1459.

[7] *Sentinel-2 user handbook*, 2nd ed., European Space Agency, Jul. 2015.

[8] "The digitalglobe constellation," Brochure, DigitalGlobe, Inc., Westminster, CO, USA, 2016.

[9] *RapidEye imagery product specifications*, 6th ed., Planet Labs Inc., Jan. 2016.

[10] L. Loncan, L. B. Almeida, J. B. Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simões, J. Y. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, and N. Yokoya, "Hyperspectral pansharpening: a review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.

[11] A. Mookambiga and V. Gomathi, "Comprehensive review on fusion techniques for spatial information enhancement in hyperspectral imagery," *Multidimensional Systems and Signal Processing*, pp. 1–17, Apr. 2016.

[12] G. Palubinskas, "Model-based view at multi-resolution image fusion methods and quality assessment measures," *International Journal of Image and Data Fusion*, pp. 203–218, May. 2016.

[13] 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.

[14] 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.

[15] 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.

[16] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "25 years of pansharpening: A critical review and new developments," in *Signal Image Processing for Remote Sensing*, 2nd ed., C. H. Chen, Ed. Boca Raton, FL: CRC Press, 7 2011, ch. 28, pp. 533–548.

[17] G. Vivone, L. Alparone, J. Chanussot, M. D. Mura, A. Garzelli, G. Licciardi, R. Restaino, and L. Wald, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.

[18] X. X. Zhu, C. Grohnfeldt, and R. Bamler, "Exploiting joint sparsity for pansharpening: The j-sparsefi algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2664–2681, May 2016.

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

[20] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," Patent US 6 011 875 A, 2000.

[21] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS + Pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.

[22] J. G. Liu, "Smoothing Filter-based Intensity Modulation: A spectral preserve image fusion technique for improving spatial details," *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3461–3472, Jan. 2000.

[23] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and Pan imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.

[24] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, Feb. 2011.

[25] X. X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, May 2013.

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

[27] 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.

[28] Z. Chen, H. Pu, B. Wang, and G.-M. Jiang, "Fusion of hyperspectral and multispectral images: A novel framework based on generalization of pan-sharpening methods," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 8, pp. 1418–1422, Aug. 2014.

[29] C. Grohnfeldt, X. X. Zhu, and R. Bamler, "Jointly sparse fusion of hyperspectral and multispectral imagery," in *Proc. IEEE IGARSS*, Jul. 2013, pp. 1–4.

[30] ——, "The J-SparseFI-HM hyperspectral resolution enhancement method Now fully automated," in *Proc. IEEE WHISPERS*, Jun. 2014, pp. 1–4.

[31] ——, "Splitting the hyperspectral-multispectral image fusion problem into weighted pan-sharpening problems – the spectral grouping concept," in *2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, June 2015, pp. 1–4.

[32] M. Selva, B. Aiazzi, F. Butera, L. Chiarantini, and S. Baronti, "Hyper-sharpening: A first approach on SIM-GA data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3008–3024, Jun. 2015.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] Q. Wei, J. M. B. Dias, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, Jul. 2015.

[39] M. Simões, J. B. Dias, L. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3373–3388, Jun. 2015.

[40] 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.

[41] 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.

[42] B. Zhukov, D. Oertel, F. Lanzl, and G. Reinhackel, "Unmixing-based multisensor multiresolution image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1212–1226, May 1999.

[43] O. Berné, A. Helens, P. Pilleri, and C. Joblin, "Non-negative matrix factorization pansharpening of hyperspectral data: An application to mid-infrared astronomy," in *Proc. IEEE WHISPERS*, Jun. 2010, pp. 1–4.

[44] N. Yokoya, T. Yairi, and A. Iwasaki, "Hyperspectral, multispectral, and panchromatic data fusion based on non-negative matrix factorization," in *Proc. IEEE WHISPERS*, Jun. 2011, pp. 1–4.

[45] ——, "Coupled non-negative matrix factorization (CNMF) for hyperspectral and multispectral data fusion: application for pasture classification," in *Proc. IEEE IGARSS*, Jul. 2011, pp. 1779–1782.

[46] R. Kawakami, J. Wright, Y. Tai, Y. Matsushita, M. Ben-Ezra, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. IEEE CVPR*, Jun. 2011, pp. 2329–2336.

[47] J. Bieniarz, D. Cerra, J. Avbelj, P. Reinartz, and R. Müller, "Hyperspectral image resolution enhancement based on specral unmixing and information fusion," in *Proc. ISPRS*, Sep. 2011, pp. 33–37.

[48] 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.

[49] E. Wycoff, T.-H. Chan, K. Jia, W.-K. Ma, and Y. Ma, "A non-negative sparse promoting algorithm for high resolution hyperspectral imaging," in *Proc. ICASSP*, May 2013, pp. 1409–1413.

[50] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *Proc. ECCV*, 2014, pp. 63–78.

[51] B. Huang, H. Song, H. Cui, J. Peng, and Z. Xu, "Spatial and spectral image fusion using sparse matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1693–1704, Mar. 2014.

[52] M. A. Bendoumi, M. He, and S. Mei, "Hyperspectral image resolution enhancement using high-resolution multispectral image based on spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6574–6583, Oct. 2014.

[53] J. Bieniarz, R. Mller, X. X. Zhu, and P. Reinartz, "Hyperspectral image resolution enhancement based on joint sparsity spectral unmixing," in *Proc. IEEE IGARSS*, 2014, pp. 2645–2648.

[54] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *Proc. IEEE ICCV*, Dec. 2015, pp. 3586–3594.

[55] ——, "Advances in hyperspectral and multispectral image fusion and spectral unmixing," in *Proc. ISPRS*, vol. XL-3/W3, 2015, pp. 451–458.

[56] M. Veganzones, M. Simões, G. Licciardi, N. Yokoya, J. Bioucas-Dias, and J. Chanussot, "Hyperspectral super-resolution of locally low rank images from complementary multisource data," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 274–288, Jan. 2016.

[57] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Bayesian fusion of multiband images," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1117–1127, Sep. 2015.

[58] N. Akhtar, F. Shafait, and A. Mian, "Bayesian sparse representation for hyperspectral image super resolution," in *Proc. IEEE CVPR*, 2015, pp. 3631–3640.

[59] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Fast fusion of multi-band images based on solving a Sylvester equation," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4109–4121, Nov. 2015.

[60] 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.

[61] 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.

[62] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proc. ICML*, 2009, pp. 1–8.

[63] J. Bioucas-Dias, "A variable splitting augmented lagrangian approach to linear spectral unmixing," in *Proc. IEEE WHISPERS*, 2009, pp. 1–4.

[64] J. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. IEEE WHISPERS*, 2010, pp. 1–4.

[65] M. V. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.

[66] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[67] M. F. Baumgardner, L. L. Biehl, and D. A. Landgrebe, "220 band AVIRIS hyperspectral image data set: June 12, 1992 Indian Pine test site 3," Sep 2015. [Online]. Available: https://purr.purdue.edu/publications/1947/1

[68] N. Yokoya and A. Iwasaki, "Airborne hyperspectral data over Chikusei," Space Application Laboratory, the University of Tokyo, Japan, Tech. Rep. SAL-2016-05-27, May 2016.

[69] C. Debes, A. Merentitis, R. Heremans, J. Hahn, N. Frangiadakis, T. van Kasteren, W. Liao, R. Bellens, A. Pižurica, S. Gautama, W. Philips, S. Prasad, Q. Du, and F. Pacifici, "Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2405–2418, Jun. 2014.

[70] E. Bedini, F. V. der Meer, and F. V. Ruitenbeek, "Use of HyMap imaging spectrometer data to map mineralogy in the Rodalquilar caldera, southeast Spain," *Int. J. Remote Sens.*, vol. 30, no. 2, pp. 327–348, 2009.

[71] K. Segl, L. Guanter, C. Rogass, T. Kuester, S. Roessner, H. Kaufmann, B. Sang, V. Mogulsky, and S. Hofer, "EeteS The EnMAP end-to-end simulation tool," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 522–530, Apr. 2012.

[72] K. Segl, L. Guanter, F. Gascon, T. Kuester, C. Rogass, and C. Mielke, "S2eteS: An end-to-end modeling tool for the simulation of Sentinel-2 image products," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5560–5571, Oct. 2015.

[73] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sensing*, vol. 63, no. 6, pp. 691–699, 1997.

[74] R. E. Roger and J. F. Arnold, "Reliably estimating the noise in AVIRIS hyperspectral images," *Int. J. Remote Sen.*, vol. 17, no. 10, pp. 1951–1962, 1996.

[75] N. Yokoya, J. C. W. Chan, and K. Segl, "Potential of resolution-enhanced hyperspectral data for mineral mapping using simulated EnMAP and Sentinel-2 images," *Remote Sensing*, vol. 8, no. 3: 172, 2016.

[76] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, J. P. Barloon, and A. F. H. Goetz, "The spectral image processing system (SIPS) - Interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, vol. 44, pp. 145–163, 1993.

[77] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?" in *Proc. the Third Conference Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images*, Jan. 2000, pp. 99–103.

[78] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[79] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.

[80] A. Garzelli and F. Nencini, "Hypercomplex quality assessment of multi/hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 662–665, Oct. 2009.

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

[82] 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, Apr. 2013.

[83] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 1995.

[84] J. J. Rodriguez and L. I. Kuncheva, "Rotation forest: A new classifier ensemble method." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1619–1630, Oct. 2006.

[85] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[86] J. Xia, P. Du, X. He, and J. Chanussot, "Hyperspectral remote sensing image classification based on rotation forest," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 239–243, Jan. 2014.

[87] C. Grohnfeldt and X. X. Zhu, "Towards a combined sparse representation and unmixing based hyperspectral resolution enhancement method," in *Proc. IEEE IGARSS*, 2015, pp. 2872–2875.

[88] R. Guerra, S. López, and R. Sarmiento, "A computationally efficient algorithm for fusing multispectral and hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5712–5728, Jun. 2016.

[89] C. Zou and Y. Xia, "Poissonian hyperspectral image superresolution using alternating direction optimization," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4464–4479, Sep. 2016.

Naoto Yokoya (S'10–M'13) received the M.Sc. and Ph.D. degrees in aerospace engineering from the University of Tokyo, Tokyo, Japan, in 2010 and 2013, respectively.

From 2012 to 2013, he was a Research Fellow with Japan Society for the Promotion of Science, Tokyo, Japan. Since 2013, he is an Assistant Professor with the University of Tokyo. Since 2015, he is also an Alexander von Humboldt Research Fellow with the German Aerospace Center (DLR), Oberpfaffenhofen, and Technical University of Munich (TUM), Munich, Germany. His research interests include image analysis and data fusion in remote sensing.

Claas Grohnfeldt (S'12) received the Diploma degree in mathematics from the University of Bremen, Germany, in 2012. During his studies, he went on exchange at the University of Warwick, England, where he studied mathematics for one year.

Since March 2012, he has been working towards the Ph.D. degree with both the DLR German Aerospace Center and the Technische Universitt Mnchen (TUM), Munich, Germany. He was awarded a Munich Aerospace doctoral research scholarship in 2012. Mr. Grohnfeldt won the first price in the DLR Science Slam competition and the IEEE IGARSS Best Paper Award in 2013 and 2016, respectively. Before his Ph.D. activities he was a research assistant at the chair Optimization and Optimal Control at the Universtiy of Bremen and a full-time intern at Airbus Defense and Space, Ottobrunn, Germany. He worked for the European Space Agency, Noordwijk, the Netherlands, and conducted his diploma thesis with Daimler AG, Stuttgart, Germany. His current research interests include signal processing in Earth observation, satellite data fusion, algorithm development and high performance computing.

Jocelyn Chanussot (M'04–SM'04–F'12) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from the Universit de Savoie, Annecy, France, in 1998. In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA - French National Defense Department). Since 1999, he has been with Grenoble INP, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing. He has been a visiting scholar at Stanford University (USA), KTH (Sweden) and NUS (Singapore). Since 2013, he is an Adjunct Professor of the University of Iceland. In 2015-2017, he is a visiting professor at the University of California, Los Angeles (UCLA).

Dr. Chanussot is the founding President of IEEE Geoscience and Remote Sensing French chapter (2007-2010) which received the 2010 IEEE GRS-S Chapter Excellence Award. He was the co-recipient of the NORSIG 2006 Best Student Paper Award, the IEEE GRSS 2011 and 2015 Symposium Best Paper Award, the IEEE GRSS 2012 Transactions Prize Paper Award and the IEEE GRSS 2013 Highest Impact Paper Award. He was a member of the IEEE Geoscience and Remote Sensing Society AdCom (2009-2010), in charge of membership development. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair (2009-2011) and Cochair of the GRS Data Fusion Technical Committee (2005-2008). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006-2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing, (2009). He was an Associate Editor for the IEEE Geoscience and Remote Sensing Letters (2005-2007) and for Pattern Recognition (2006-2008). Since 2007, he is an Associate Editor for the IEEE Transactions on Geoscience and Remote Sensing. He was the Editor-in-Chief of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2011-2015). In 2013, he was a Guest Editor for the Proceedings of the IEEE and in 2014 a Guest Editor for the IEEE Signal Processing Magazine. He is a Fellow of the IEEE and a member of the Institut Universitaire de France (2012-2017).