Shadow-Aware Nonlinear Spectral Unmixing With Spatial Regularization

Guichen Zhang, Paul Scheunders, *Senior Member, IEEE*, and Daniele Cerra, *Member, IEEE*

***Abstract*— Current shadow-aware hyperspectral unmixing (HySU) methods often suffer from noisy abundance maps and inaccurate abundance estimation of shadowed pixels, as these are characterized by low reflectance values and signal-to-noise ratio. In order to achieve a shadow-insensitive abundance estimation, in this article, we propose a novel spatial–spectral shadow-aware mixing (S3AM) model. The approach models shadows by considering diffuse solar illumination and secondary illumination from neighboring pixels. Besides, spatial regularization using shadow-aware weighted total variation (TV) is employed. Specifically, pixels in the local neighborhood of a target pixel simultaneously consider spectral similarity measures derived from the imagery, elevation similarity measures derived from a digital surface model (DSM), and the impact of shadows. The sky view factor F, needed as input for the model, is also derived from available DSMs. The proposed approach is extensively validated and compared with state-of-the-art methods on two datasets. Results demonstrate that the S3AM yields superior abundance estimation maps for real scenarios, by decreasing the noise in the results and achieving more accurate reconstructions in the presence of shadows.**

***Index Terms*— Digital surface model (DSM), shadow aware, spatial regularization, spectral mixing model, spectral unmixing, total variation (TV).**

I. Introduction

SPECTRAL unmixing is a fundamental hyperspectral image analysis technique analyzing the composition of an image element at subpixel level [1], [2], [3]. Given a spectrum related to an image element, spectral unmixing decomposes it into the spectra related to the pure materials, which are present (usually more than one), i.e., endmembers, and their corresponding contributions, i.e., abundances [4]. In supervised unmixing, the endmembers present in the scene are available after being extracted directly from the image using endmember extraction algorithms, or collected in an external spectral library [1]. Subsequently, only abundances need to be estimated. In this article, we exclusively discuss supervised unmixing.

Manuscript received 18 March 2023; revised 10 June 2023; accepted 19 June 2023. Date of publication 26 June 2023; date of current version 7 August 2023. *(Corresponding author: Guichen Zhang.)*

Guichen Zhang is with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Weßling, Germany, and also with the Department of Computer Science, Osnabrück University, 49074 Osnabrück, Germany (e-mail: guichen.zhang@outlook.com).

Paul Scheunders is with the Imec-Vision Lab, Department of Physics, University of Antwerp, 2610 Antwerp, Belgium (e-mail: paul.scheunders@uantwerpen.be).

Daniele Cerra is with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Weßling, Germany (e-mail: daniele.cerra@dlr.de).

Digital Object Identifier 10.1109/TGRS.2023.3289570

A spectral mixing model describes how an incoming light ray from a given illumination source interacts with the targets on ground, before it is scattered back to the spectrometer [5]. One of the most important prerequisites to correctly carry out the spectral unmixing process is to have a spectral mixing model at hand that properly represents the scenario [2]. In the past decades, numerous spectral mixing models with different physical assumptions have been proposed [6], [7], [8], [9], [10], [11]. The linear mixing model (LMM) [6] is one of the most popular approaches, as it achieves a good balance between simplicity and accuracy in the physical modeling of the scattering process. Specifically, the LMM assumes that incoming light interacts with ground materials only once before being scattered back to the sensor. In recent years, non-LMMs have attracted attention, as nonlinear effects in optical interactions are occasionally nontrivial and meaningful for a correct interpretation of the resulting spectra [1], [2]. At a macroscopic level, nonlinear effects often appear due to height differences between ground objects, causing incoming light rays to interact more than once with different objects before being scattered back to the sensor. Nonlinear algorithms model secondary optical reflections of a light ray using the termwise product of the spectra related to the objects on its path. Most methods regard nonlinearities up to the second order, including the Nascimento model [7], the Fan model [8], the polynomial post-nonlinear model (PPNM) [11], and the generalized bilinear model (GBM) [9]. Recently, Heylen and Scheunders [10] presented the multilinear mixing (MLM) model to tackle nonlinear effects of all orders of optical interaction based on stochastic processes.

Another challenge for spectral unmixing is posed by shadow effects. A given pixel can be fully or partly sunlit or shadowed, with partial shadowing mainly related to shadow boundaries. Most existing spectral mixing models assume all pixels to be sunlit [1], [2]. However, different illumination conditions apply for shadowed pixels [12]. Specifically, fully sunlit pixels receive both direct and diffuse solar radiation, i.e., global solar illumination, while fully shadowed pixels only receive diffuse solar radiation, caused by atmospheric scattering [12]. For those spectral mixing models that do not consider shadow effects [6], [7], [8], [9], [10], [11], the difference in solar illumination conditions between sunlit and shadowed regions can lead to inaccurate abundance estimation, as shadows considerably impact the magnitude of the reflectance and introduce significant wavelength-dependent spectral distortion [12]. Therefore, advanced spectral mixing models that consider

shadow effects are desirable. Unless otherwise stated, we will refer in this manuscript to shadowed pixels as fully or partly shadowed image elements, while sunlit pixels denote fully sunlit image elements.

Several methods have been developed to address this issue. Since shadowed pixels have lower reflectance values, one of the most straightforward ways to describe a shadowed pixel is by scaling its measured spectrum. This is motivated by the fact that a fully shadowed region does not receive direct light from any light source, while fully sunlit areas receive global solar illumination. In addition, partly shadowed areas can be considered to be a mixture of fully sunlit and shadowed pixels. Earlier works include an additional "shade" endmember whose spectral values are (nearly) zero, followed by abundance estimation using the linear spectral unmixing model [13], [14]. Equivalently, the shadow LMM (SLMM) [5] extends the LMM by subtracting the "shade" fraction from a pixel using a parameter Q. In [5], the shadow fraction parameter Q has been applied in the MLM model to jointly consider nonlinear and shadow effects, resulting in the shadow MLM (SMLM) model. This group of methods then relies on the shadow-related parameter Q to remove shadows. Assuming that shadows scale the spectral reflectance, the spectral angle is a shadow-insensitive measure of spectral similarities. The strategy in [15] is to conduct nonlinear spectral unmixing separately in sunlit and shadowed regions, followed by matching endmembers in shadowed areas with those in sunlit areas using the spectral angle distance (SAD). The shadow removal is conducted using pixel reconstruction while replacing shadowed endmembers with their matched sunlit ones.

However, the above approaches ignore an essential illumination contribution in shadowed regions, namely, the diffuse solar illumination, which comes from the optical scattering of the direct sunlight in the atmosphere [12]. This contribution has a wavelength-dependent impact on shadows, which is nonnegligible and nontrivial for spectral modeling [12]. One straightforward way of dealing with this wavelength-dependent shadow effect is to develop physics-based mixing models based on radiative transfer [12]. Specifically, those methods include two major solar illumination sources, i.e., direct and diffuse solar radiation, with any pixel receiving a combination of both [16], [17], [18], [19]. For example, Uezato et al. [16] proposed an illumination-invariant spectral mixing model at radiance level based on radiative transfer. In [17], the Fansky model was proposed, that accounts for the shadow effects by incorporating direct and diffuse solar radiation, and that accounts for nonlinear optical interactions by using the Fan model [8]. A shadow-compensated bilinear mixing model (SCBMM) in [18] allows for different illumination conditions in a scene and estimates the diffuse solar radiation and abundances simultaneously based on global particle swarm optimization [20]. Zhang et al. [19] introduced the extended shadow-aware MLM (ESMLM) model, describing the spectrum of a mixed pixel as the sum of the contributions from different illumination sources.

Nevertheless, several shadow-related challenges remain. Since the contribution from diffuse illumination is significantly smaller with respect to global illumination, modeling shadows with diffuse solar illumination terms may lead to overfitting of the optimization problem. Moreover, the signal-to-noise ratio in shadowed pixels is much lower with respect to sunlit pixels. These factors lead to inaccurate abundance estimations and noisy abundance maps [19].

The abovementioned problems may be solved by exploiting the spatial information in shadow-aware spectral unmixing methods. Spatial–spectral unmixing approaches consider the spatial dependence in local and nonlocal neighborhoods. On the one hand, in local neighborhoods, abundances at a specific pixel are assumed to be strongly correlated with the ones of neighboring pixels [21], [22]; on the other hand, in nonlocal neighborhoods, similar patches in a larger region are assumed to share similar texture [23], [24]. Earlier work [25] improved the local homogeneity of abundance maps by iteratively reducing the spatial structure of residual maps. Later, a hierarchical Bayesian model incorporated spatial information based on Markov random fields for spectral unmixing [26]. Yao et al. [27] account for the spatial–contextual information through the convolutional operation and sequentially decompose the hyperspectral image from local attention to global aggregation. Recently, convolutional neural networks (CNNs) have been incorporated into deep-learning-based spectral unmixing approaches [28] for the modeling of spatial information. Moreover, a group of methods applies spatial regularization to spectral mixing models. In the past decades, numerous works have applied different spatial regularizers to abundances, such as spatial autocorrelation [29], spatial–spectral coherence [30], $L2$ norm [31], total variation (TV) [32], nonlocal TV [23], and nonlocal HSI TV incorporating the spatial distribution of the endmembers [24]. Specifically, the TV regularization has attracted most attention, as it promotes piecewise smooth abundance maps and better preserves edges [32], [33].

While TV assumes neighboring pixels to have an equal influence to a target pixel, recently some works refined this assumption by allowing different weighting factors at neighboring pixels, resulting in weighted TV [33], [34], [35]. Neighboring pixels with associated bigger weights are assumed to have a larger impact on the underlying spectrum. Weights can be derived from different measures as prior knowledge, such as spectral distance, principal components, and abundance distance [34], [35]. In addition to spectral information, weights derived from elevation information can improve abundance estimation considerably [33] because of the illumination-insensitive characteristics of digital surface models (DSMs). In general, the more accurately, the weights describe the ground features, the better the abundance estimation performance. On the other hand, inaccurate weights can lead to imprecise results. For example, since shadows introduce wavelength-dependent spectral distortions, spectral similarity measures can largely reduce performance at shadows boundaries, where pixels composed of the same target material may exhibit spectral differences when exposed to different illumination conditions.

In order to partly resolve the above challenges, in this article, we propose a spectral unmixing method with shadow-aware spatial constraints obtained from a hyperspectral image

and a corresponding DSM generated by multiview stereo imagery.

1) Inspired by our previous work in [19], the spectral mixing model accounts for typical ground scenarios in the presence of shadows and nonlinear optical effects by considering multiple illumination sources: global solar illumination, diffuse solar illumination, and secondary illumination from neighboring pixels. Specifically, global solar radiation is assumed to be the main illumination source for sunlit pixels, while diffuse solar radiation dominates in shadowed pixels. A ground pixel may receive light from some or all of the illumination sources, creating flexible spectral modeling for pixels with different illumination conditions.
2) In order to alleviate the impact of shadows on the estimation of abundances, we propose a weighted TV constraint with shadow-insensitive weighting factors. Weighting factors are computed from the spectral angle and elevation differences between a target pixel and its neighboring pixels. In addition, a precomputed shadow-related parameter is included in weighting factors in order to decrease the contribution of shadowed neighboring image elements of the target pixel.
3) We inject elevation information from the DSM into the model in two ways. First, the elevation data provide illumination-insensitive TV weights, beneficial to the abundance estimation in shadowed pixels. Second, rather than being an additional model parameter, the sky view factor (F), required to calculate the contribution of diffuse solar illumination, is previously obtained from the elevation data, additionally decreasing the complexity of the spectral mixing model.
4) We extensively validate the proposed method on two real hyperspectral images, including shadows, both quantitatively and qualitatively. The proposed model significantly decreases the noise level in abundance maps, shows good robustness to shadow effects, and obtains more homogeneous abundance maps.

The remainder of this article is organized as follows. In Section II, we present the most popular spectral mixing model, i.e., LMM, and the ESMLM model, a shadow-aware spectral unmixing approach from our earlier works [19]. Section III introduces the proposed spatial–spectral shadow-aware mixing (S3AM) model with spatial constraints. Section IV describes the experimental setup, including datasets, compared methods, measures for quantitative evaluation, optimal parameter settings, and computational resources. Sections V and VI report and assess experimental results on two real datasets. Finally, we conclude our work and give prospects for future developments in Section VII.

II. Spectral Mixing Models

We first introduce the mathematical notations utilized in this article. Denote a hyperspectral image with B spectral bands and N pixels as $\boldsymbol{X} \in \mathbb{R}^{B\times N}$, with $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N]$, where pixel $\boldsymbol{x}_j = (x_{j,1}, x_{j,2}, \ldots, x_{j,B})^T \in \mathbb{R}^{B\times 1}$. An endmember library $\boldsymbol{E} \in \mathbb{R}^{B\times p}$ consists of p endmembers, where the ith endmember is denoted as $\boldsymbol{e}_i \in \mathbb{R}^{B\times 1}$. The averaged spectrum in the first-order neighborhood of pixel $\boldsymbol{x}_j$ is denoted as $\boldsymbol{\chi}_j \in \mathbb{R}^{B\times 1}$. Denote an abundance matrix related to $\boldsymbol{E}$ as $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_N]$, with $\boldsymbol{a}_j \in \mathbb{R}^{p\times 1}$. In addition, four pixelwise parameters at pixel j are denotes as P_j, Q_j, K_j, and F_j, and their corresponding vector forms are denoted as $\boldsymbol{P}$, $\boldsymbol{Q}$, $\boldsymbol{K}$, and $\boldsymbol{F}$.

Spectral mixing models can be constructed by a ray-based description of the interaction of the incoming light with the ground materials [5], [19]. Following some physical assumptions, a light ray initiated from an illumination source interacts with ground materials with given probabilities before being scattered back to the sensor. In a simple scenario, we assume that all pixels receive only global solar illumination, i.e., direct sunlight, and that light rays interact only once with the ground materials before being scattered back to the sensor. Moreover, the probability of the optical interaction within a single-resolution cell is assumed to be proportional to the abundance of each material within the image element. Following the above assumptions, the LMM is obtained as follows:

$$\boldsymbol{x}_j = \sum_{i=1}^{p} a_{j,i}\boldsymbol{e}_i \tag{1}$$

where $\sum_{i=1}^{p} a_{j,i} = 1$ and $\forall i\colon a_{j,i} \geq 0$.

The ESMLM model described in [19] is based on similar ray-based descriptions. In order to account for shadows, the ESMLM model allows various illumination conditions in an image and accounts for typical scenarios related to the types and distribution of ground materials. Specifically, the ESMLM model considers three illumination sources: global solar illumination, diffuse solar illumination, and secondary illumination from neighboring pixels. A light ray from each illumination source follows certain physical assumptions. For a given pixel $\boldsymbol{x}_j$, the ESMLM model sums up contributions of possible light rays initiated from all three illumination sources, as follows:

$$\begin{aligned}\boldsymbol{x}_j = {} & (1-Q_j)(1-P_j)\sum_{i=1}^{p} a_{j,i}\boldsymbol{e}_i + Q_j T(s_{0_{\text{diff}}}) \odot \sum_{i=1}^{p} a_{j,i}\boldsymbol{e}_i \\ & + a_{j,i}\boldsymbol{e}_i P_j \sum_{i_1=1}^{p}\sum_{i_2=1}^{p} a_{j,i_1}a_{j,i_2}\boldsymbol{e}_{i_1} \odot \boldsymbol{e}_{i_2} \\ & + (1-Q_j)(1-P_j)K_j \sum_{i=1}^{p} a_{j,i}\boldsymbol{e}_i \odot \boldsymbol{\chi}_j \end{aligned} \tag{2}$$

where $\sum_{i=1}^{p} a_{j,i} = 1$, $\forall i\colon a_{j,i} \geq 0$, and $P_j, Q_j, K_j \in [0, 1]$.

We briefly remind the main concepts related to the ESMLM model and its representation reported in (2). For further details, the interested reader is referred to the extensive description in [19].

In this model, P is the probability that a light ray will undergo a further interaction with the current pixel before reaching the sensor, while Q denotes the shadowed fraction of a pixel. The first term in (2) represents the linear optical interactions of the incoming light from global solar illumination. This term is rescaled with $(1 - Q)$, since the shadowed part of a pixel does not have a direct line of sight

to the sun. Moreover, it is rescaled with $(1-P)$, representing the probability that it goes directly to the sensor after this interaction. The second term is related to the linear interactions of incoming light from diffuse solar illumination, contributing to the shadowed fraction Q of a pixel. The operator $T(s_{0_{\text{diff}}})$ corresponding to the diffuse solar illumination is modeled in (3). Furthermore, the third term describes the second-order optical reflections, appearing with probability P. The fourth term describes secondary interactions between the first-order neighboring pixels (described by spectrum $\boldsymbol{\chi}_j$) and the target pixel with a strength factor K. By only retaining nonlinear effects up to the second order, only neighboring regions with a direct line of sight to the sun can contribute to the target pixel, resulting in a rescaling factor of $(1-Q)(1-P)$. Furthermore

$$T(s_{0_{\text{diff}}}) = \frac{\boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{E}_s}{\boldsymbol{\tau}_{\text{dir}} \odot \boldsymbol{E}_l + \boldsymbol{\tau}_{\text{diff}} \odot \boldsymbol{E}_s} \tag{3}$$

where $(\tau_{\text{diff}}(\lambda)E_s(\lambda)/\tau_{\text{dir}}(\lambda)E_l(\lambda)) = F(k_1\lambda^{-k_2} + k_3)$ with $k_1, k_2, k_3 > 0$. This power function models the wavelength-dependent atmospheric scattering, which is stronger at shorter wavelengths. $\boldsymbol{E}_l$ and $\boldsymbol{E}_s$ represent direct and diffuse solar radiation, respectively, while $\boldsymbol{\tau}_{\text{dir}}$ and $\boldsymbol{\tau}_{\text{diff}}$ represent the transmittance of direct and diffuse solar radiation, respectively. If a ground pixel is not occluded, the diffuse radiation comes from all directions of the sky. When occlusion occurs, the diffuse illumination decreases by the sky view factor $F \in [0, 1]$, representing the fraction of sky that a ground pixel can "see."

The ESMLM model provides flexible nonlinear modeling with four parameters (P, Q, K, and F) and accounts for different illumination conditions in an image element. Such flexibility brings challenges in solving the reverse problem due to the nonconvexity of the objective function. In particular, the ESMLM model becomes tri-convex, making it rather challenging to acquire a satisfying solution through the alternating direction method of multiplier (ADMM) approach [36], [37].

III. Proposed Method

In this article, we propose an S3AM model by embedding of spatial information. In order to make our problem biconvex for an improved convergence [36], [37], we simplified the ESMLM model in two aspects. First, we set $P = 0$, because the inner-pixel second-order optical interactions have been observed to have minor impact on spectral unmixing results. Besides, we assume that the neighboring pixels contribute equally to a target pixel regardless of their illumination conditions, so that the neighbor illumination term is rescaled solely according to parameter K.

The simplified model at pixel j is given by

$$\begin{aligned}\boldsymbol{x}_j &= (1-Q_j)\boldsymbol{y}_j + Q_j\boldsymbol{y}_j \odot \widetilde{\boldsymbol{f}}_j + K_j\boldsymbol{y}_j \odot \boldsymbol{\chi}_j \\ &= (\mathbf{1}_B - \mathbf{1}_B Q_j) \odot \boldsymbol{y}_j + \mathbf{1}_B Q_j \odot \boldsymbol{y}_j \odot \widetilde{\boldsymbol{f}}_j + \mathbf{1}_B K_j \boldsymbol{y}_j \odot \boldsymbol{\chi}_j \\ &= \boldsymbol{E} \odot (\mathbf{1}_B - \mathbf{1}_B Q_j + \mathbf{1}_B Q_j \odot \widetilde{\boldsymbol{f}}_j + \mathbf{1}_B K_j \odot \boldsymbol{\chi}_j)\mathbf{1}_p^T \boldsymbol{a}_j \\ &= \widetilde{\boldsymbol{E}}_j \boldsymbol{a}_j \end{aligned} \tag{4}$$

where

$$\widetilde{\boldsymbol{f}}_j = \frac{F_j \cdot (k_1\boldsymbol{\lambda}^{-k_2} + k_3)}{1 + F_j \cdot (k_1\boldsymbol{\lambda}^{-k_2} + k_3)} \tag{5}$$

$$\boldsymbol{y}_j = \boldsymbol{E}\boldsymbol{a}_j = \sum_{i=1}^{p} a_{j,i}\boldsymbol{e}_i. \tag{6}$$

$$\widetilde{\boldsymbol{E}}_j = \boldsymbol{E} \odot \left((\mathbf{1}_B - \mathbf{1}_B Q_j + \widetilde{\boldsymbol{f}}_j Q_j + \boldsymbol{\chi}_j K_j)\mathbf{1}_p^T\right). \tag{7}$$

We construct the optimization problem in vector form as follows:

$$\min_{\boldsymbol{a}_j, Q_j, K_j} \frac{1}{2}\sum_{j=1}^{N} \|\widetilde{\boldsymbol{E}}_j \boldsymbol{a}_j - \boldsymbol{x}_j\|_F^2. \tag{8}$$

The abundance nonnegativity constraint (ANC) and abundance sum-to-one constraint (ASC) are applied on the abundances $\boldsymbol{a}_j$ [5], [19]. In addition, we assume Q and $K \in [0, 1]$, in order to maintain their physical meanings

$$\boldsymbol{a}_j \geq 0, \sum_{i=1}^{p} \boldsymbol{a}_{j,i} = 1, \; Q, \quad K \in [0, 1]. \tag{9}$$

Inspired by the existing works on weighted TV constraints for spectral unmixing [33], [34], the following spatial constraint on the abundances is proposed:

$$\sum_{j=1}^{N} \sum_{m \in \mathcal{N}(j)} R_{j,m} \|\boldsymbol{a}_j - \boldsymbol{a}_m\|_1^1 \tag{10}$$

where $\mathcal{N}(j)$ denotes the first-order neighborhood of the target pixel j. $R_{j,m}$ represents a weighting factor describing the similarity between pixel j and m

$$R_{j,m} = \frac{1}{Z_j}\left(R_{h_{j,m}} + R_{x_{j,m}}\right) \tag{11}$$

where $R_{h_{j,m}}$ and $R_{x_{j,m}}$ represent the weighting factors related to height and spectral information, respectively. Z_j is the normalizing constant value and constraints the summation of weighting factors at pixel j to 1.

$R_{h_{j,m}}$ is defined as follows:

$$R_{h_{j,m}} = \exp\left[-\frac{1}{\delta_h^2}(1 + \eta Q'_{j,m})T_{h_{j,m}}\right] \tag{12}$$

where δ_h^2 is a constant parameter controlling the weight range and η reduces the influence of shadowed neighboring pixels on a target pixel using $Q'_{j,m}$, the shadow fraction at the neighboring pixel m for the target pixel j. In this article, $Q'_{j,m}$ is precalculated using the SLMM method [5]. Finally, $T_{h_{j,m}}$ is a height similarity measure, given by the normalized height difference between pixels j and m

$$T_{h_{j,m}} = \frac{(h_j - h_m)^2}{(h_j + h_m)^2} \tag{13}$$

normalized surface height above the ellipsoid h is provided by the DSM, which is illumination-insensitive and, therefore, robust to shadow effects. Thus, neighboring pixels with larger height similarities will have larger impact on the target pixel.

The weighting factor $R_{x_{j,m}}$ corresponds to spectral information

$$R_{x_{j,m}} = \exp\left[-\frac{1}{\delta_x^2}\left(1 + \eta Q'_{j,m}\right)T_{x_{j,m}}\right] \tag{14}$$

where δ_x^2 is a constant parameter controlling the weight range, the shadow-related parameters η and $Q'_{j,m}$ are the same as

in (12), and the spectral similarity measure $T_{x_{j,m}}$ is defined by the spectral angle [38], [39]

$$T_{x_{j,m}} = \max\left(\arccos\frac{\boldsymbol{x}_j \cdot \boldsymbol{x}_m}{\|\boldsymbol{x}_j\|\|\boldsymbol{x}_m\|} - 0.1, 0\right). \tag{15}$$

Since shadow effects introduce spectral distortions [12], [40], the spectral angle between sunlit and shadowed pixels of the same material can be significantly larger than 0. We found this difference empirically to be around 0.1. In order to mitigate the impact of distortion in the spectral similarity measure, a value of 0.1 is then subtracted from the spectral angle in (15) up to a minimum value of 0.

Furthermore, as nonlinear effects typically do not depend on spectral, height, and shadow conditions, we apply a non-weighed TV constraint on K

$$\sum_{j=1}^{N}\sum_{m\in\mathcal{N}(j)}\|K_j - K_m\|_1^1. \tag{16}$$

Hence, we define the optimization problem with spectral and spatial constraints as follows:

$$\min_{\boldsymbol{A},\boldsymbol{Q},\boldsymbol{K}}\frac{1}{2}\sum_{j=1}^{N}\|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j - \boldsymbol{x}_j\|_F^2 + \lambda\|\boldsymbol{A}\boldsymbol{W}_1\|_{1,1} + \ell_{\mathcal{C}}(\boldsymbol{A}) + \ell_{\mathcal{S}}(\boldsymbol{A}) + \ell_{\mathcal{M}}(\boldsymbol{Q}) + \lambda\|\boldsymbol{K}\boldsymbol{W}_2\|_{1,1} + \ell_{\mathcal{M}}(\boldsymbol{K}) \tag{17}$$

where $\ell_{\mathcal{C}}(\boldsymbol{A}) = \{\boldsymbol{A}|\boldsymbol{A} \geq \boldsymbol{0}_{p\times N}\}$, $\ell_{\mathcal{S}}(\boldsymbol{A}) = \{\boldsymbol{A}|\boldsymbol{1}_p^T\boldsymbol{A} = \boldsymbol{1}_N^T\}$, $\ell_{\mathcal{M}}(\boldsymbol{Q}) = \{\boldsymbol{Q}|\boldsymbol{Q} \geq \boldsymbol{0}_{1\times N}, \boldsymbol{Q} \leq \boldsymbol{1}_{1\times N}\}$, and $\ell_{\mathcal{M}}(\boldsymbol{K}) = \{\boldsymbol{K}|\boldsymbol{K} \geq \boldsymbol{0}_{1\times N}, \boldsymbol{K} \leq \boldsymbol{1}_{1\times N}\}$. The sparse matrix $\boldsymbol{W}_2 = [\boldsymbol{W}_2^{\uparrow}\boldsymbol{W}_2^{\downarrow}\boldsymbol{W}_2^{\leftarrow}\boldsymbol{W}_2^{\rightarrow}] \in \mathbb{R}^{N\times 4N}$, where each element belongs to the set $\{-1, 0, 1\}$, consists of differential operators in four directions, i.e., up, down, left, and right. $\boldsymbol{K}\boldsymbol{W}_2$ computes the difference in K in each direction in the first-order neighborhood of each pixel. For instance, the difference in K in the upward direction at pixel j can be written as $K_{m^{\uparrow}} - K_j$, where $m^{\uparrow}$ denotes the index of the neighboring pixel in the upward direction of pixel j. Similarly, the sparse matrix $\boldsymbol{W}_1 = [\boldsymbol{W}_1^{\uparrow}\boldsymbol{W}_1^{\downarrow}\boldsymbol{W}_1^{\leftarrow}\boldsymbol{W}_1^{\rightarrow}] \in \mathbb{R}^{N\times 4N}$ consists of differential operators in four directions weighted by the factor $R_{j,m}$. $\boldsymbol{A}\boldsymbol{W}_1$ computes the difference in A in each direction in the first-order neighborhood of each pixel for each endmember, weighted by the factor $R_{j,m}$. For instance, the abundance difference in the upward direction at pixel j associated with endmember i can be written as $(a_{m^{\uparrow},i} - a_{j,i})R_{j,m^{\uparrow}}$, where $m^{\uparrow}$ denotes the index of the neighboring pixel in the upward direction of pixel j.

The above optimization is a biconvex problem, and it is convex to $\boldsymbol{A}$ and $\{\boldsymbol{Q}, \boldsymbol{K}\}$, respectively. Following [36] and [37], we split the unknown variables into two groups and solve two convex problems sequentially using the ADMM approach. In the ADMM form, the optimization problem is given by

$$\min_{\boldsymbol{A},\boldsymbol{Q},\boldsymbol{K},\boldsymbol{G},\boldsymbol{H}}\frac{1}{2}\sum_{j=1}^{N}\|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j - \boldsymbol{x}_j\|_F^2 + \lambda\|\boldsymbol{G}_2\|_{1,1} + \ell_{\mathcal{C}}(\boldsymbol{G}_3) + \ell_{\mathcal{S}}(\boldsymbol{G}_4) + \ell_{\mathcal{M}}(\boldsymbol{H}_1) + \lambda\|\boldsymbol{H}_3\|_{1,1} + \ell_{\mathcal{M}}(\boldsymbol{H}_4)$$
$$\times\begin{cases}\boldsymbol{G}_1 = \boldsymbol{A}\\ \boldsymbol{G}_2 = \boldsymbol{G}_1\boldsymbol{W}_1\\ \boldsymbol{G}_3 = \boldsymbol{A}\\ \boldsymbol{G}_4 = \boldsymbol{A}\end{cases}\begin{cases}\boldsymbol{H}_1 = \boldsymbol{Q}\\ \boldsymbol{H}_2 = \boldsymbol{K}\\ \boldsymbol{H}_3 = \boldsymbol{H}_2\boldsymbol{W}_2\\ \boldsymbol{H}_4 = \boldsymbol{K}.\end{cases} \tag{18}$$

The updating sequence is reported in Algorithm 1, while updating equations of primal and dual variables are derived in the Appendix.

Algorithm 1 ADMM for the Optimization Problem of (18)

Input : $\boldsymbol{E}$, $\widetilde{f}$, $\boldsymbol{X}$, $\boldsymbol{\chi}$, $\boldsymbol{\lambda}$, k_1, k_2, k_3
Output: $\boldsymbol{A}$, $\boldsymbol{Q}$, $\boldsymbol{K}$
Initialize: $t = 0$, $\boldsymbol{A}^{(0)}$, $\boldsymbol{Q}^{(0)}$, $\boldsymbol{K}^{(0)}$, $\boldsymbol{G}^{(0)}$, $\boldsymbol{H}^{(0)}$, $\boldsymbol{U}^{(0)}$
while *the stopping criterion is not satisfied* **do**
  Given $\boldsymbol{Q}^{(t)}$, $\boldsymbol{K}^{(t)}$, $\boldsymbol{G}^{(t)}$, $\boldsymbol{U}^{(t)}$, update $\boldsymbol{A}^{(t+1)}$ with (24)
  Given $\boldsymbol{A}^{(t+1)}$, $\boldsymbol{U}^{(t)}$, update $\boldsymbol{G}^{(t+1)}$ with (25), (26), (27), (28)
  Given $\boldsymbol{A}^{(t+1)}$, $\boldsymbol{H}^{(t)}$, $\boldsymbol{U}^{(t)}$, update $\boldsymbol{Q}^{(t+1)}$ and $\boldsymbol{K}^{(t+1)}$ with (30), (31), (32)
  Given $\boldsymbol{Q}^{(t+1)}$, $\boldsymbol{K}^{(t+1)}$, $\boldsymbol{U}^{(t)}$, update $\boldsymbol{H}^{(t+1)}$ with (33), (34), (35), (36)
  Given $\boldsymbol{U}^{(t)}$, $\boldsymbol{A}^{(t+1)}$, $\boldsymbol{Q}^{(t+1)}$, $\boldsymbol{K}^{(t+1)}$, $\boldsymbol{G}^{(t+1)}$, $\boldsymbol{H}^{(t+1)}$, update $\boldsymbol{U}^{(t+1)}$ with (37)
  t = t + 1
end

IV. Experimental Setup

We validate our proposed method on real data, relying on the German Aerospace Center (DLR) HyperSpectral Unmixing (HySU) and the HySpex/4K datasets. The DLR HySU benchmark dataset is used for quantitative and qualitative validation. It contains hyperspectral imagery with reference abundances available for ground materials. The original dataset does not contain shadows and a DSM, so we additionally simulated both in this article for validation purposes. Besides, in order to demonstrate our proposed method on real shadows and a real DSM, we employ the HySpex and 4K datasets. As the abundances in the latter dataset are unknown, we solely carry out a qualitative validation in this case.

A. Datasets

1) DLR HySU Benchmark Dataset: The image in Fig. 1(a) was acquired over Oberphaffenhofen, Bavaria, Germany with a HySpex push-broom camera, resulting in a ground sampling distance of 0.7 m. The image comprises 135 spectral bands ranging from 417.4 to 902.8 nm. This dataset [41] contains five square ground targets with a side of 3 m (bitumen, red-painted metal sheets, blue fabric, red fabric, and green fabric) and a background material (grass), for a total of six known endmembers [Fig. 1(c)]. The fully constrained linear spectral unmixing method [42] is applied to the shadow-free image, yielding pixelwise reference abundance maps. When using the ground target sizes to evaluate abundance estimation errors [41], the reference abundance maps have an average error of 2.3% [41], which is low enough to justify their use as ground truth in our experiments.

A binary shadow map is drawn in order to shade a part of all targets, followed by a Gaussian filter with a size of 3×3 to create a soft shadow mask Q [Fig. 1(b)]. Furthermore,

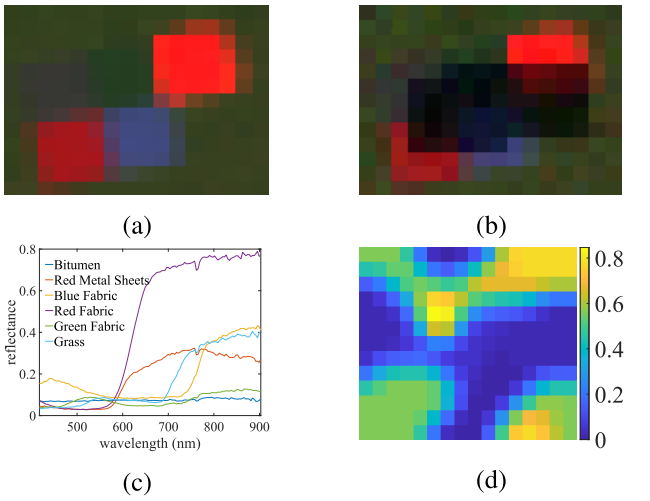

Fig. 1. DLR HySU dataset with simulated shadowed pixels. (a) Hyperspectral image as a true-color composite including five ground targets with side lengths of 3 m. (b) Hyperspectral image with simulated shadowed pixels and additional random noise (SNR = 30). (c) Endmember library containing five targets (bitumen, red-painted metal sheets, blue fabric, red fabric, and green fabric) and surrounding grass. (d) Simulated DSM.

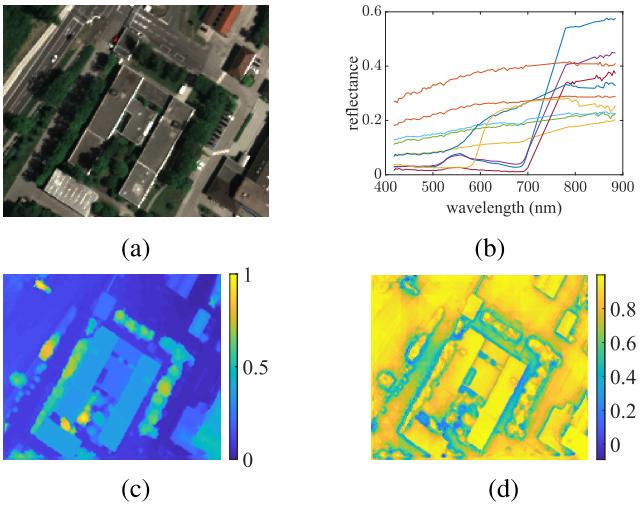

Fig. 2. Hyperspectral dataset. (a) Hyperspectral image as a true-color composite acquired by the HySpex sensor in the study area of Oberpfaffenhofen, Bavaria, Germany. (b) Endmember library, automatically extracted from (a). (c) Normalized DSM. (d) Sky view factor map derived from the DSM in (c).

a synthetic height map is simulated through a piecewise homogeneous distribution using the Potts model [26], followed by Gaussian filtering. Then, the synthetic height map is used to compute the simulated sky view factor F, using the method in [43]. Given the shadow-free pixel $\boldsymbol{y}_j$, we simulate the pixel $\boldsymbol{x}_j$ with artificial shadows as follows:

$$\boldsymbol{x}_j = (1 - Q_j)\boldsymbol{y}_j + Q_j T(s_{0_{\text{diff}}})\boldsymbol{y}_j \tag{19}$$

where atmospheric parameters k_1, k_2, and k_3 are set as detailed in Section IV-D. Furthermore, we apply additional random noise (SNR = 30) on the simulated image.

2) HySpex and 4K Dataset: This dataset consists of an airborne hyperspectral image and a DSM, acquired at the same time over Oberpfaffenhofen, Bavaria, Germany between 8:42 and 8:56 A.M. [Central European Summer Time (CEST)] on June 4, 2018. The airborne hyperspectral image [Fig. 2(a)] was acquired with a HySpex visible and near-infrared (VNIR) sensor and has been atmospherically corrected using atmospheric and topographic correction (ATCOR) [41], [44], [45]. After removing the water vapor bands, a total of 101 bands have been kept for further processing. Moreover, endmembers $\boldsymbol{E}$ have been extracted from the fully sunlit pixels using the method in [19] and [46] [see Fig. 2(b)]. Multiview stereo imagery acquired with the 4K camera system was employed to generate the DSM [47] in Fig. 2(c), whose values represent surface height above the ellipsoid. In addition, height values were normalized within $[0, 1]$ to retain the relative height of the ground surface. After geometrical co-registration and resampling, the DSM and images share the same geocoordinates and spatial resolution (i.e., 0.7 m). Given the height data, the sky view factor F was computed using the software System for Automated Geoscientific Analyses (SAGA) [43] [see Fig. 2(d)].

B. Methods in Comparison

We compare the S3AM method with the following state-of-the-art spectral mixing models.

1) *LMM [6]:* An LMM that does not take shadows into account.
2) *SLMM [13]:* An LMM accounting for shadows using a scaling factor while ignoring diffuse solar illumination.
3) *SMLM [5]:* A non-LMM accounting for shadows using a scaling factor as in SLMM, along with nonlinear interactions using the MLM model [10].
4) *NUEM [15]:* A shadow removal method based on nonlinear unmixing and endmember matching. This method requires a shadow mask as input, which has been computed by S3AM in the experiment.
5) *Fansky [17]:* A non-LMM considering shadows based on both direct and diffuse solar illumination, along with nonlinear interactions using the Fan model [8].
6) *ESMLM [19]:* An extended SMLM model incorporating direct and diffuse solar illumination as well as nonlinear interactions.
7) *SCBMM [18]:* A non-LMM considering shadows using both direct and diffuse solar illumination and nonlinear interactions using the bilinear mixing model (BMM) [7].

Moreover, we investigate the impact of different variations of the spatial TV regularization in an ablation study.

C. Quantitative Measures

Given the observed pixels $\boldsymbol{x}_j$ and the reconstructed pixels $\hat{\boldsymbol{x}}_j$, the mean reconstruction error (RE) is defined as follows:

$$\text{RE} = \frac{1}{N}\sum_{j=1}^{N}\sqrt{\sum_{\lambda=1}^{B}(x_{j,\lambda} - \hat{x}_{j,\lambda})^2} \tag{20}$$

while the mean abundance error (AE) is defined as follows:

$$\text{AE} = \frac{1}{pn}\sum_{j=1}^{N}\sum_{i=1}^{p}|a_{j,i} - \hat{a}_{j,i}|. \tag{21}$$

D. Parameter Settings

1) k_1, k_2, and k_3: Following our previous work in [19], k_1, k_2, and k_3 are the parameters of a power function that

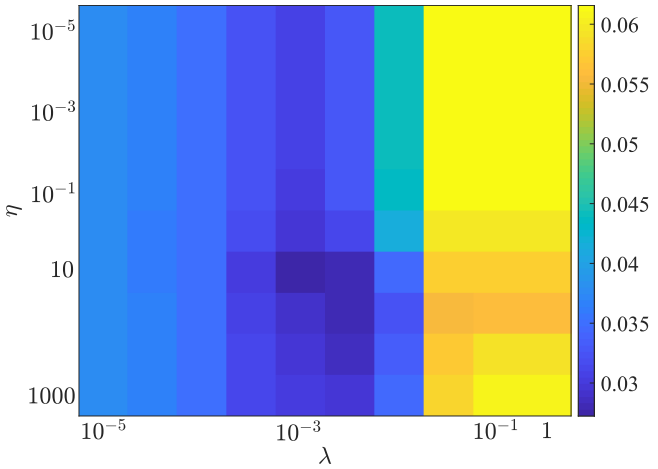

Fig. 3. AE as a function of λ and η for the DLR HySU dataset, where $\lambda \in \{10^{-5}, 5 \cdot 10^{-5}, 10^{-4}, 5 \cdot 10^{-4}, 10^{-3}, 5 \cdot 10^{-3}, 0.01, 0.05, 0.1, 1\}$ and $\eta \in \{10^{-5}, 10^{-4}, 10^{-3}, 0.01, 0.1, 1, 10, 100, 500, 1000\}$.

models the ratio of diffuse to global solar irradiance on the ground surface. We assume that atmospheric conditions are consistent in the entire image region, so these parameters are constant in our experiments. In practice, we compute these parameters using ten pairs of pixels that are selected in the scene. Specifically, a pair of pixels contain a sunlit pixel and a fully shadowed pixel near a shadow boundary, assumed to consist of the same material. We avoid including vegetation during the selection in order to avoid dealing with complex nonlinear effects. Then, k_1, k_2, and k_3 are solved by (4), with $K_j = 0$, $Q_j = 1$, and $p = 1$. The obtained values are follows: $k_1 = 0.579$, $k_2 = 6.974$, and $k_3 = 0.206$.

2) λ and η: Fig. 3 presents the AE as a function of λ and η for the DLR HySU dataset. We calculate optimal values of these parameters from the values $\lambda \in \{10^{-5}, 5 \cdot 10^{-5}, 10^{-4}, 5 \cdot 10^{-4}, 10^{-3}, 5 \cdot 10^{-3}, 0.01, 0.05, 0.1, 1\}$ and $\eta \in \{10^{-5}, 10^{-4}, 10^{-3}, 0.01, 0.1, 1, 10, 100, 500, 1000\}$ by minimizing AE, resulting in $\lambda = 10^{-3}$ and $\eta = 10$ for S3AM. As ground-truth abundances are not available for the HySpex dataset, we empirically determine λ and η to be the same as for the DLR HySU dataset. In addition, we analyze the impact of different values of λ on the abundance maps in Section VI.

3) δ_x^2 and δ_h^2: These parameters represent the weight range in the exponential functions of the height-related (12) and spectral (14) weighting factors, respectively. In principle, one can optimize the values of δ_x^2 and δ_h^2 in a similar way as λ and η, by minimizing the optimization error. However, too many free parameters can lead to overfitting. In practice, more than 99.9% of T_x and T_h was found to lie within the range $[0, 0.5]$. Hence, we set the weighting ranges within the same span and choose empirical values $\delta_x^2 = \delta_h^2 = 0.1$.

4) μ: The penalty parameter μ of the augmented Lagrangian [see (23)] was determined as in [32]. The initial value is set to $\mu = 0.001$ and is then updated iteratively by keeping the ratio between primal and dual residual norms within a positive value of 10, as suggested in [36].

5) Initialization and Stopping Criteria: A fully constrained spectral unmixing method [42] based on the SLMM model has been applied to initialize $\boldsymbol{A}$ and $\boldsymbol{Q}$, while $\boldsymbol{K}$ and $\boldsymbol{U}$ are initialized to zero. In addition, the algorithm stops when the primal residual is less than $5 \cdot 10^{-4}$ or the maximum number of iterations, set as 100, is reached.

TABLE I
RUNNING TIME OF COMPARED METHODS

Method	Running time (s)		
	HySU (208 pixel)	subset1 (1148 pixel)	subset 2 (1085 pixel)
LMM	0.47	2.56	2.27
SLMM	0.61	5.80	5.38
SMLM	0.70	9.42	8.33
NUEM	0.63	6.01	4.65
Fansky	7.42	112.75	109.20
ESMLM	2.59	26.66	25.47
SCBMM	11.43	42.25	40.43
S3AM	0.55	2.80	2.71

E. Computational Resources

All algorithms were developed in a MATLAB environment and run on an Intel Core i7 −8650 U CPU, 1.90-GHz machine with four cores and eight logical processors. We apply the MATLAB function FMINCON using the sequential quadratic programming algorithm to perform the (non)linear optimization for LMM, SLMM, SMLM, nonlinear unmixing and endmember matching (NUEM), Fansky, and ESMLM. The function and constraint tolerance are set to 10^{-6} and 10^{-5}, respectively. In addition, for-loop iterations over all pixels were running in parallel on workers in a parallel pool.

Table I shows the running time of compared methods. The Fansky and SCBMM have the highest computational cost, depending on the number of pixels in the subset. Besides, the nonlinear optimization implemented by FMINCON in ESMLM, SMLM, and SLMM requires more computational resources with respect to the linear optimization in LMM. The S3AM, solved by the ADMM approach, has a short computation time, as MATLAB efficiently computes closed-form updates for this method.

V. RESULTS: DLR HYSU BENCHMARK DATASET

A. Pixel Reconstruction and Abundance Estimation

Fig. 4 shows the obtained AE and RE for all methods, from sunlit, (partly) shadowed, and all pixels, respectively. The AE was obtained without taking grass into account. The reason for this is that reference abundance values were derived by LMM. In grass areas, nonlinear effects may be present, and this class contains nonnegligible intraclass variations with respect to the spectrum selected as endmember. Therefore, reference abundances for grass may be unreliable. In addition, NUEM is not included in the comparison of RE, as this method runs two unmixing processes followed by spectral matching. Furthermore, we show AE and RE on a degraded image with additive noise (SNR = 30), in order to evaluate the robustness to noise of the compared methods.

All compared methods show satisfactory REs in sunlit regions. In shadowed regions, the LMM obtains significantly higher RE compared with other methods, indicating the importance of shadow-aware modeling. Nevertheless, smaller REs do not necessarily imply a satisfactory abundance estimation. Compared with REs, we observe significantly larger differences of AEs among compared methods. In general, the better a model accounts for shadows, the better the

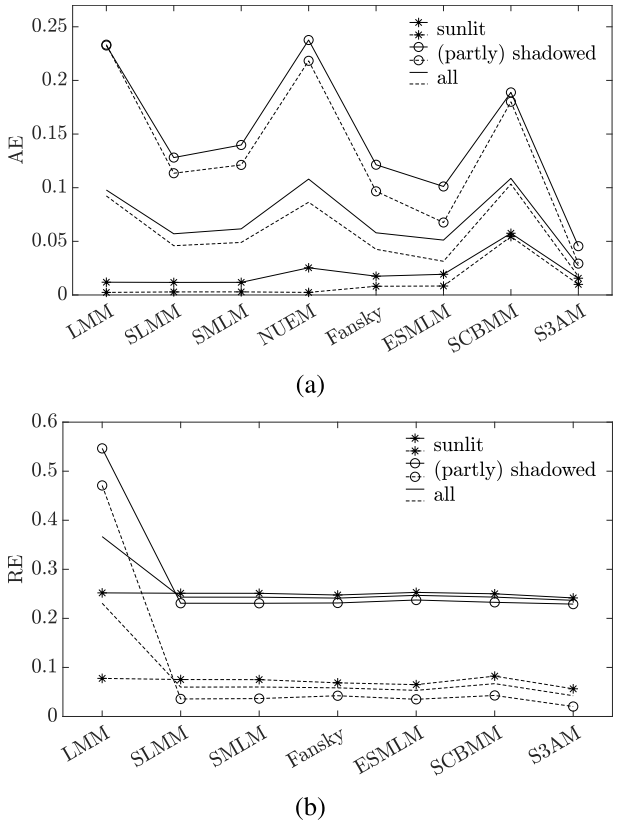

Fig. 4. Comparison of (a) mean AE and (b) mean RE for the DLR HySU dataset. Solid and dashed lines represent results obtained using the input image with and without additional noise (SNR = 30), respectively. Sunlit and (partly) shadowed pixels are determined with $Q \leq 0.1$ and $Q > 0.1$, respectively.

abundance estimation. The LMM has the lowest performance, because shadow effects are simply ignored. The NUEM is also characterized by a large AE. Since shadow effects cause a wavelength-dependent spectral distortion, it is quite challenging to perform spectral matching between sunlit and shadowed pixels. Moreover, the SLMM and SMLM treat shadows as a scaling effect and perform better at abundance estimation in shadowed regions. The performance improves further when including the diffuse solar illumination as prior knowledge. Fansky, ESMLM, and S3AM select pairs of pixels from the input image as prior knowledge and estimate the ratio of diffuse and global solar illumination through a power function. The proposed approach, i.e., S3AM, clearly outperforms others because of the applied spatial constraints. In addition, the ratio of diffuse and global solar illuminations can be estimated along with abundance values in the unmixing process, i.e., SCBMM. Nevertheless, its AE appears higher than other methods. The reason is that the SCBMM does not assume stronger atmospheric scattering at shorter wavelengths [12]. Hence, its estimated diffuse radiation may not correspond to the spectral characteristics of shadows in practice. Furthermore, results in Fig. 4 show that both REs and AEs increase considerably, as the image is degraded by noise. Despite better abundance estimation, the Fansky and ESMLM are less robust to noise compared with LMM, SLMM, SMLM, NUEM, and SCBMM. Since the contribution of diffuse solar radiation is significantly smaller relative to global radiation, the optimization problem can be overfit and may lead to noisy abundance maps. This sensitivity to noise can be significantly alleviated by the proposed method S3AM because of the spatial constraints.

A qualitative comparison of the abundance maps is displayed in Fig. 5. In order to assess the location of AEs, we overlay the AE map in gray scale as a semitransparent layer over the optical image in Fig. 6. First of all, it can be observed that AEs are mainly located in shadowed regions. Compared with the LMM, the shadow-aware unmixing methods therein display improved abundance maps. While SLMM and SMLM treat shadow as a scaling effect, the SMLM further considers nonlinear reflections. Since the study region is a flat terrain with artificial materials, multiple reflections appear minor. Thus, the abundance maps of SLMM and SMLM are very similar. Unlike embedding a shadow-related parameter in the model, the NUEM matches spectra in shadowed areas with their corresponding ones in sunlit regions. However, accurate spectral matching is rather challenging due to spectral distortion caused by shadows. In particular, this dataset contains some materials with similar spectral information, making spectral matching more difficult. For example, the NUEM easily confuses red-painted metal sheets and red fabric.

Furthermore, we compare methods that take into account diffuse solar illumination, i.e., SCBMM, Fansky, ESMLM, and S3AM. The SCBMM may have the ability to estimate abundance values in shadowed regions, such as bitumen, while it seems challenging to estimate diffuse solar illumination without prior knowledge. Fansky, ESMLM, and S3AM, on the other hand, whose diffuse solar illumination is computed from manually selected pixels in the input image, perform better in abundance estimation.

However, without applying spatial constraints, abundance maps show higher noise levels and confusion between similar materials, such as red-painted metal sheets and red fabric. Such distortions can be alleviated by injecting spatial information into the analysis. By applying weighted TV constraints, S3AM considerably improves the abundance estimation step. First, the noise level has been significantly reduced owing to the spatial constraints. In addition, the abundance estimation at the boundary pixels is significantly improved, due to the weighting of the spectral and height information in the spatial constraints.

B. Ablation Study

The S3AM method consists of a weighted TV constraint term, where the weights are formed by spectral and height features. In the ablation study, we investigate the individual contribution from each feature.

We refer to the weighted TV (10) in the S3AM method as wTVfull, where both height and spectral features are included and computed by (12) and (14). In the ablation study, we regard height and spectral features one at a time by setting $R_{h_{j,m}} = 0$ and $R_{x_{j,m}} = 0$ in (11), respectively, resulting in the ablated TV forms wTVhei and wTVspec. In addition, we set the weights $R_{h_{j,m}} = R_{x_{j,m}} = 1$ in (11), resulting in a classic nonweighted TV, labeled as TV. Moreover, we ablate the spatial constraints entirely by setting $\lambda = 0$ in (17), reducing the method to only the spectral mixing model, labeled as "none."

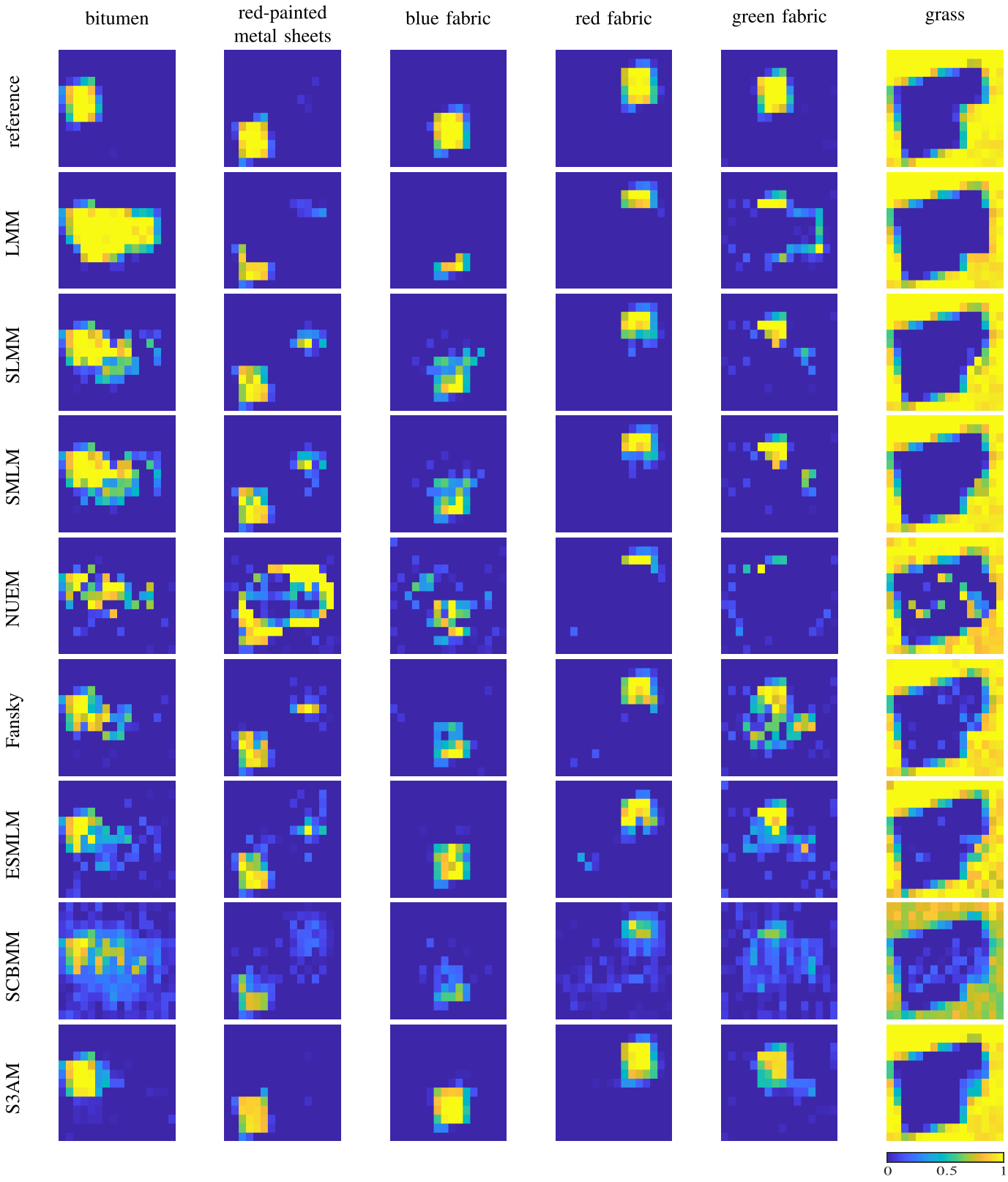

Fig. 5. Abundance maps for the HySU dataset. Left to right: bitumen, red metal sheets, blue fabric, red fabric, green fabric, and grass. Top to bottom: reference, LMM, SLMM, SMLM, NUEM, Fansky, ESMLM, SCBMM, and S3AM. The reference abundance maps are computed by applying fully constrained least squares unmixing using the library of known endmembers on the shadow-free image.

Fig. 7 compares AE of wTVfull, wTVspec, wTVhei, TV, and "none" in the ablation study. Specifically, we investigate AE in sunlit, (partly) shadowed, and all regions. Compared with sunlit regions, where the spatial constraints play a minor role, we observe considerable improvement in shadowed pixels by embedding spatial constraints, with wTVfull achieving the best abundance estimation, both with and without additional noise.

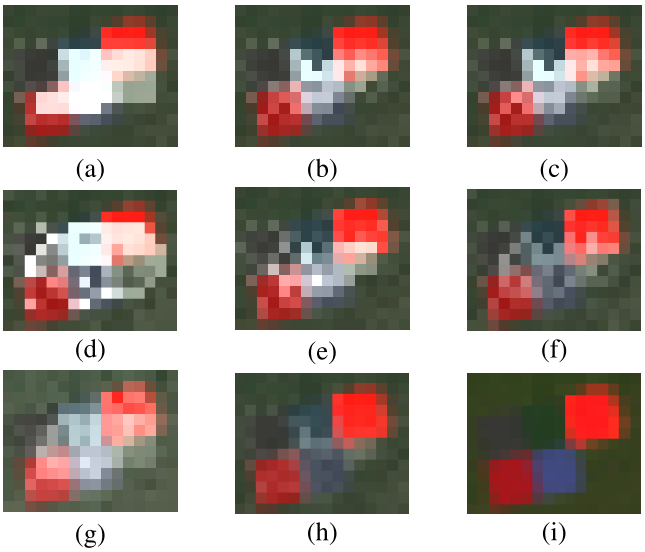

Fig. 6. DLR HySU imagery overlaid with pixelwise mean AE maps in gray scale. (a) LMM. (b) SLMM. (c) SMLM. (d) NUEM. (e) Fansky. (f) ESMLM. (g) SCBMM. (h) S3AM. (i) Reference image.

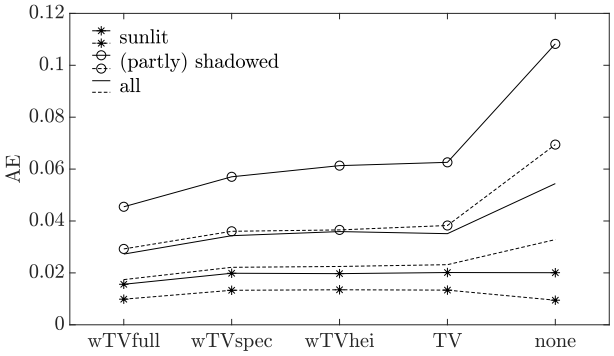

Fig. 7. Comparison of mean AE of S3AM with ablated spatial constraints for the DLR HySU dataset. Solid and dashed lines represent results obtained on the input image with and without additional noise (SNR = 30), respectively.

Figs. 8 and 9 show, respectively, RE and AE maps. When no spatial constraints are applied, i.e., "none," resulting abundance maps are noisy. The ablation study shows that differences in abundances mainly appear on boundaries between different materials. Typically, TV oversmooths the boundaries between different materials, since it treats neighboring pixels equally. One example is visible at the transition from bitumen to green fabric. When applying weighted TV, the abundance estimation on the boundary pixels is largely improved, and wTVspec better preserves the shape of the bitumen target with respect to wTVhei, because spectral information can better separate the two materials on the boundary between them. Since, in our experiment, the DSM is randomly generated, the height information does not fully correspond to the ground objects, leading to inaccurate abundance estimations. In practice, spectral or height information alone might not be sufficient to distinguish ground materials. The configuration denoted as wTVfull, adopted by the proposed method S3AM, jointly considers spectral and height features and outperforms single-source weights both quantitatively and qualitatively. Specifically, wTVfull constraints visibly improve the abundance estimation on the boundary pixels, e.g., at the right border of bitumen and the bottom side of green fabric.

C. Shadow-Removed Pixel Reconstruction

Spectral mixing models considering shadows allow generating shadow-removed imagery through pixel reconstruction. Specifically, in SLMM, SMLM, ESMLM, and S3AM, the parameter Q represents the shadow fraction within a pixel. By setting $Q = 0$, shadows are removed during pixel reconstruction [19]. The NUEM, Fansky, and SCBMM methods compute abundance values separately in sunlit and shadowed groups. Then, shadow-removed pixels can be reconstructed using summed-up abundance values in the two groups and sunlit endmember spectra. Obviously, a better abundance estimation leads to better reconstruction and shadow removal. Fig. 10 compares the shadow-removed images computed by the different models. Despite confusion between bitumen and green fabric, the SLMM achieves a satisfactory restoration in shadowed areas, which is remarkable considering its simplicity. The SMLM shows a similar abundance estimation as the SLMM (see Fig. 5). However, some dark pixels appear in the reconstructed image of the SMLM because of its incorrect estimation of parameters P and Q [17]. Large areas of red-painted metal sheets appear in the restored image of NUEM because of the mismatch between endmembers extracted in sunlit and shadowed regions. For example, when minimizing the spectral angle, the spectrum of shadowed red fabric is associated with the red-painted metal sheets. In addition, the spectrum of shadowed green fabric is associated with bitumen. Moreover, Fansky, ESMLM, and SCBMM show material confusion between different red materials. The material boundaries in shadowed areas also appear reddish due to incorrect abundance estimation. Overall, the proposed method S3AM achieves the best qualitative shadow-removed image because of its superior abundance estimation and spatial constraints.

VI. Results: HySpex and 4K Datasets

A. Abundance Estimation

Unlike for the DLR HySU dataset, we do not have ground-truth abundances for the HySpex/4K dataset. Therefore, in this section, we investigate and compare the methods qualitatively on two image subsets (Figs. 11 and 12). Fig. 11 compares the abundances of the roof material. The LMM seems to perform satisfactorily in shadowed pixels. However, it is worth noticing that the LMM easily overestimates impervious surfaces, as can be observed in subset 2, where many vegetation pixels are incorrectly recognized as impervious surfaces [Fig. 12(c)]. Besides, the SLMM, SMLM, and NUEM show lower abundances in shadowed pixels on the roof [Fig. 11(d)–(f)]. Specifically, they confuse the spectra of the roof with other impervious materials (see Section VI-C), since these two materials contain similar spectral information. Therefore, in order to distinguish between similar spectra in the presence of shadows, it is essential to consider the diffuse solar radiation [see Figs. 11(g)–(j)]. In contrast, when determining materials with large spectral differences, such as vegetation and road in subset 2 (Fig. 12), SLMM and SMLM can also achieve satisfactory results. Moreover, by considering the diffuse solar illumination, Fansky, ESMLM, and SCBMM may achieve

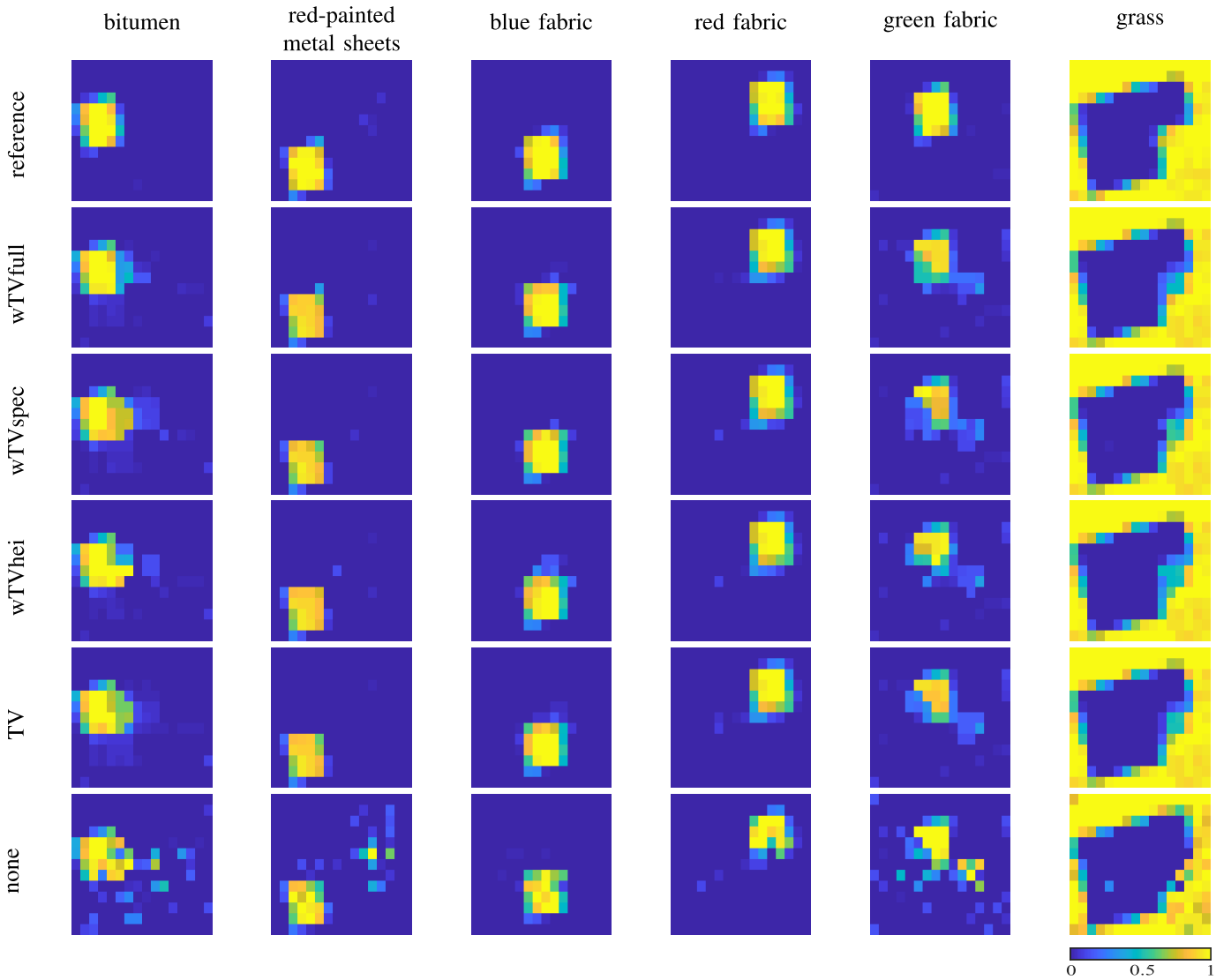

Fig. 8. Abundance maps for the HySU dataset using S3AM with ablated spatial constraints. Left to right: bitumen, red metal sheets, blue fabric, red fabric, green fabric, and grass. Top to bottom: reference, wTVfull, wTVspec, wTVhei, TV, and "none." The reference abundance maps are computed by applying fully constrained least squares using the library of known endmembers on the shadow-free image.

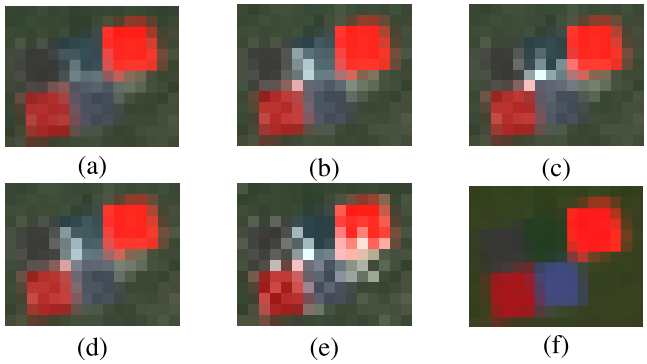

Fig. 9. DLR HySU imagery overlaid with pixelwise mean AE maps in grayscale using S3AM with ablated spatial constraints. (a) wTVfull. (b) wTVspec. (c) wTVhei. (d) TV. (e) "None." (f) Reference image.

better abundance estimation at higher noise levels. In particular, the SCBMM shows a quite noisy abundance map in subset 1. Similar to the DLR HySU dataset, the SCBMM may show decreased performance at some sunlit pixels [Fig. 12(i)]. In addition, the Fansky method appears noisier compared with ESMLM and can only distinguish a part of the shadowed materials. Compared with Fansky and SCBMM, the ESMLM performs consistently better in both subsets.

The TV constraint further contributes to the abundance estimation in two aspects. First, given the large spectral variability in real hyperspectral imagery, the spectral unmixing methods without the TV constraint can easily confuse similar materials, thus producing considerably noisier abundance maps, while the spatial constraint promotes similar abundances in local neighborhoods, significantly reducing noise. Second, the abundance estimation is not as accurate in shadowed regions, where pixels contain a lower signal-to-noise ratio. The spatial constraint provides additional information to spectral models, thus achieving better abundance estimations.

An ablation study has been conducted in subset 1, using spatial constraints wTVfull, wTVspec, wTVhei, TV, and "none" in S3AM (see Fig. 13). As in the case of the HySU dataset, differences in TV constraints mainly affect mixed sunlit/shadowed pixels in the HySpex dataset. Specifically, the weighted TV methods, i.e., wTVspec, wTVhei, and wTVfull,

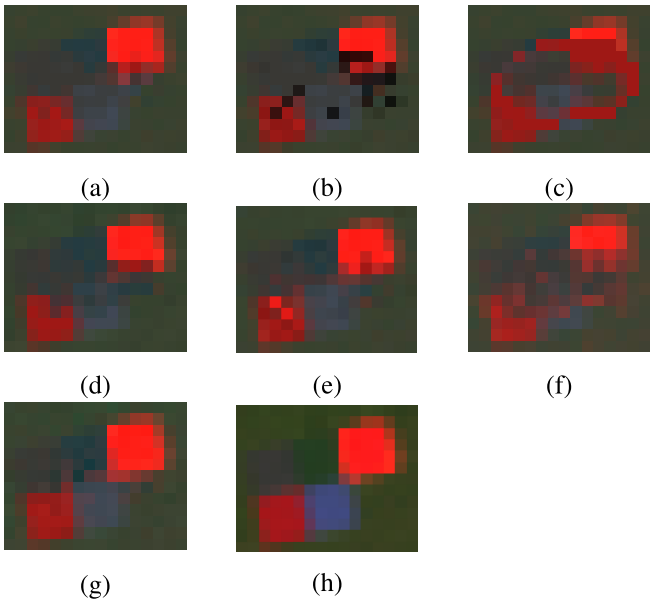

Fig. 10. Restored images with removed shadows of the DLR HySU dataset generated by (a) SLMM, (b) SMLM, (c) NUEM, (d) Fansky, (e) ESMLM, (f) SCBMM, (g) S3AM, and (h) reference image.

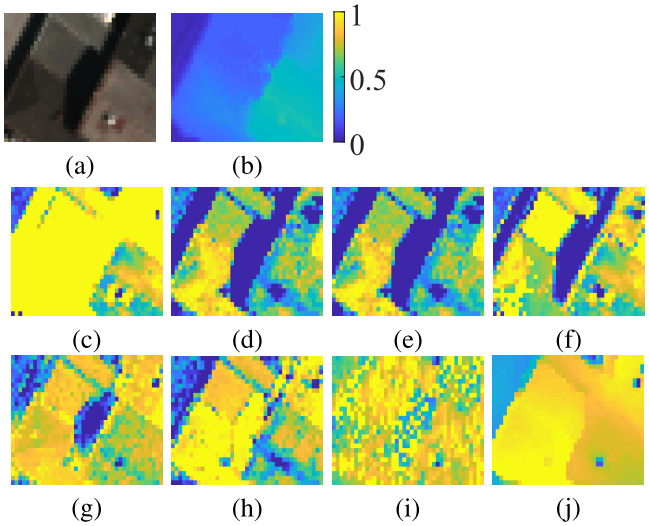

Fig. 11. Subset 1 of the HySpex dataset—(a) True-color composite and (b) DSM. Abundance maps of roof using (c) LMM, (d) SLMM, (e) SMLM, (f) NUEM, (g) Fansky, (h) ESMLM, (i) SCBMM, and (j) S3AM.

lead to sharp edges, while the classic TV method oversmooths transitions in boundary pixels. Examples can be spotted in the regions between the two roofs in subset 1.

B. Impact of λ on Abundance Estimation

Since it is very challenging to select an optimal λ for the HySpex dataset, due to the lack of ground truth, the sensitivity of various TV constraints with respect to λ is evaluated in Fig. 14. In our experiments so far, we used the optimized values of λ from the DLR HySU dataset on the HySpex dataset. However, a similar dataset with ground truth may be not available at all in real-case scenarios. Fig. 14 compares abundances of roof in subset 1 for $\lambda \in \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. As expected, the abundance maps become more homogeneous and less noisy, as λ increases. The abundance maps using $\lambda = 10^{-3}$ present the best qualitative results. As discussed, the

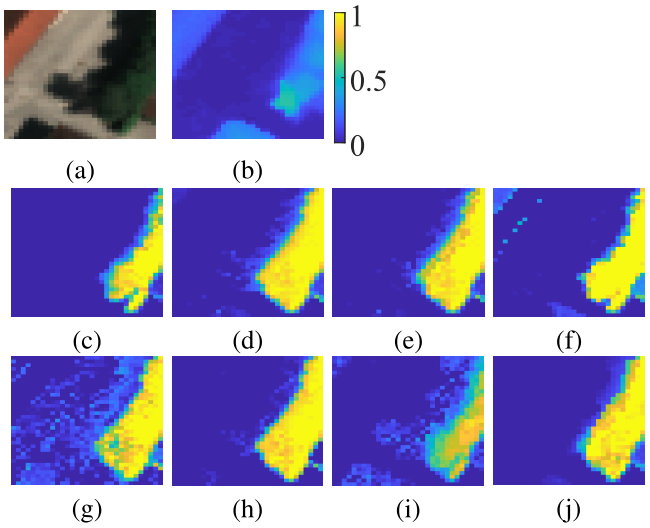

Fig. 12. Subset 2 of the HySpex dataset—(a) true-color composite and (b) DSM. Abundance maps of vegetation using (c) LMM, (d) SLMM, (e) SMLM, (f) NUEM, (g) Fansky, (h) ESMLM, (i) SCBMM, and (j) S3AM.

use of the weighted TV leads to better abundance estimation on boundary pixels, where the wTVfull achieves a balance between height and spectral information. This characteristic becomes more prominent with larger values of λ. Specifically, the TV constraint using spectral information only, i.e., wTVspec, largely decreases the performance when $\lambda \geq 10^{-2}$, as the computed weights do not match well with the ground materials. Compared with TV and wTVspec, the wTVhei constraint leads to better results, but it can suffer from height inaccuracies. Despite some boundary pixels possibly being affected for larger values of λ, the wTVfull constraint generally reaches a more robust abundance estimation in a broader range of values of λ.

C. Shadow-Removed Pixel Reconstruction

Fig. 15 shows three examples of shadow removal results for the HySpex dataset. The restored images appear very bright in shadowed roof pixels when applying the SLMM, SMLM, and NUEM approaches. This corresponds to the poor abundance estimations in Fig. 11(d)–(f). In addition, some dark pixels appear in the restored image by the SMLM because of the incorrect estimation of the parameters P and Q. A similar artifact can be observed in the DLR HySU dataset [Fig. 10(b)]. Furthermore, it is worth noticing that, when SAD successfully matches sunlit and shadowed pixels in a scene, the NUEM can achieve good results, such as in the second and third examples. However, this may work only in simple scenarios. In addition, the NUEM highly depends on the input shadow mask. Some artifacts appear when shadowed areas are over- or underestimated (see examples 1 and 3 in Fig. 15). Approaches considering the diffuse solar illumination achieve in all cases better shadow-removed images, while the results of SCBMM are much noisier. Despite the improvement, shadow removal by Fansky and ESMLM in the first example shows some spectral distortions. Finally, the reconstructed road computed by ESMLM and SCBMM in the third example appears noisy, with the line marking resulting almost invisible.

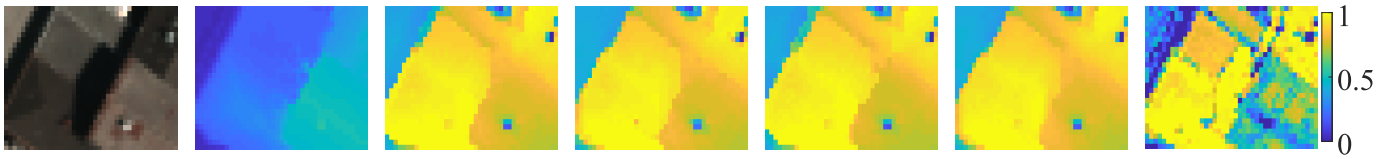

Fig. 13. Subset 1 of the HySpex dataset—(a) true-color composite and (b) DSM. Abundance maps of roof for ablation study using spatial constraints—(c) wTVfull, (d) wTVspec, (e) wTVhei, (f) TV, and (g) "none."

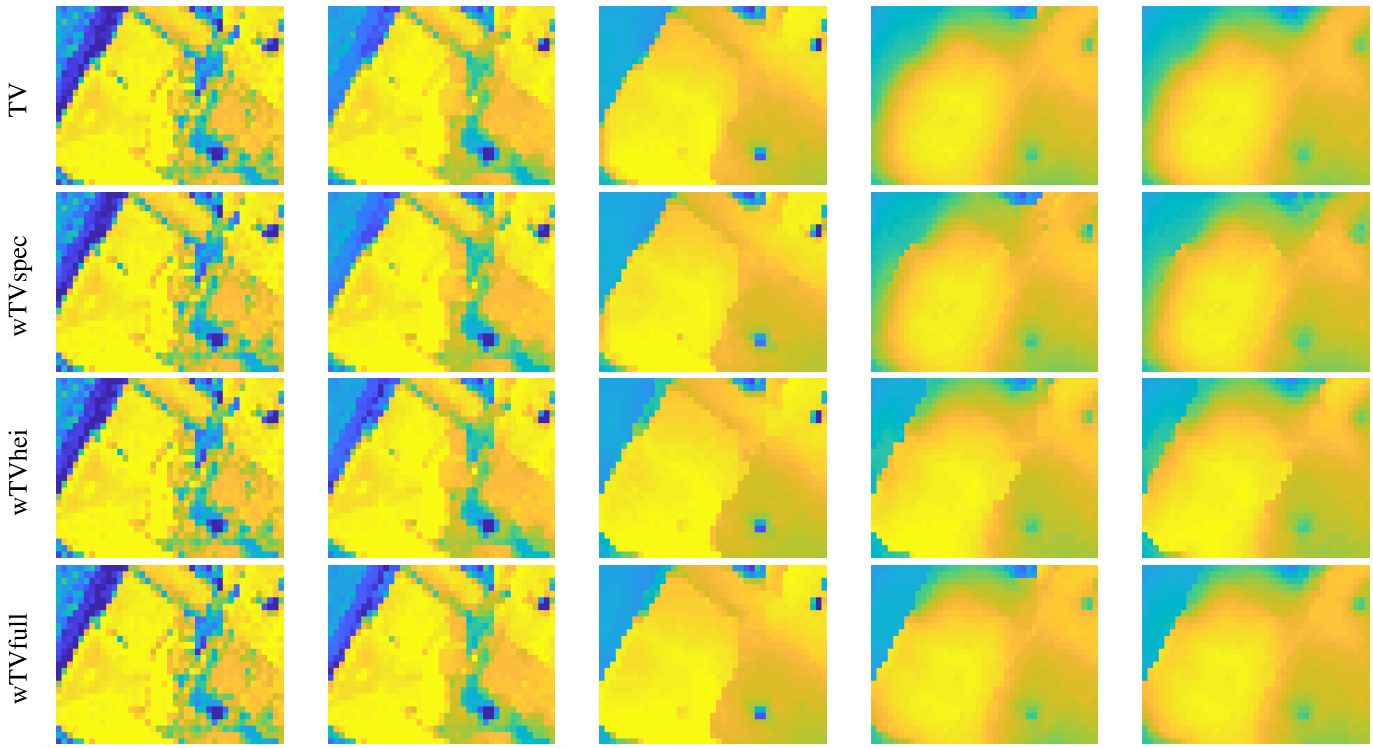

Fig. 14. Abundance maps of roof as a function of λ for subset 1 of the HySpex dataset using S3AM with ablated spatial constraints. Top to bottom: TV, wTVspec, wTVhei, and wTVfull. Left to right: $\lambda = 10^{-5}$, 10^{-4}, 10^{-3}, 10^{-2}, and 10^{-1}.

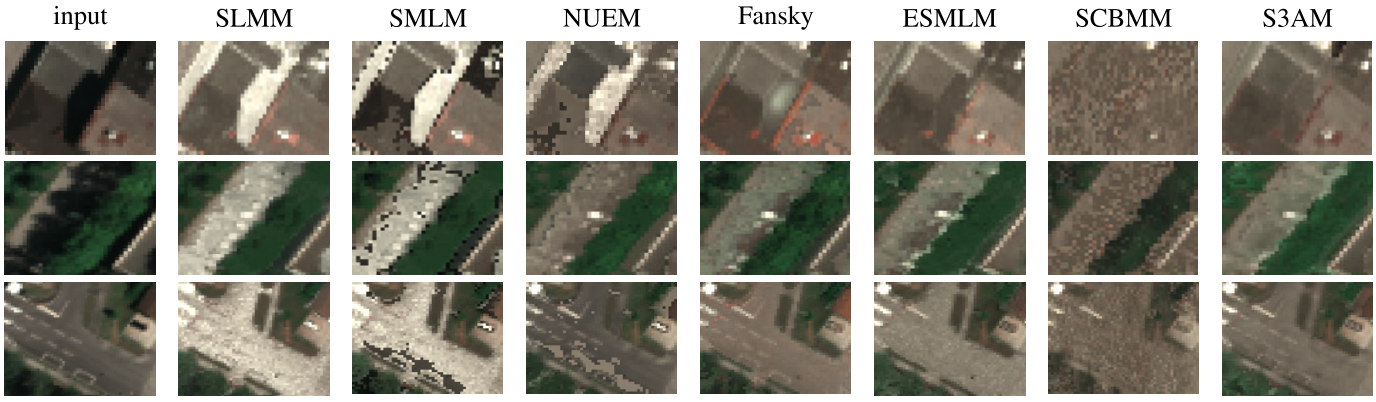

Fig. 15. Three examples of shadow-removed results for the HySpex dataset. Left to right: input image, SLMM, SMLM, NUEM, Fansky, ESMLM, SCBMM, and S3AM.

The advantages of S3AM compared with other methods appear evident. First, S3AM improves the restoration result in shadowed regions and retains spectral homogeneity. The shaded roof area in the first example of Fig. 15 shows that S3AM better reconstructs spectral features compared with ESMLM. Moreover, S3AM considerably reduces the noise in the shadow-removed imagery. This noise reduction can be spotted not only in shadowed regions, such as regions shaded by vegetation in the second example, but also in sunlit pixels, such as the impervious surface in the third example.

VII. Conclusion

This article proposes an S3AM model. The spectral modeling accounts for shadows, following physical assumptions based on radiative transfer theory. Specifically, a light path initiates from an illumination source and interacts with

endmembers before being scattered back to the observer. The model considers direct, diffuse, and neighboring illumination sources, where direct solar radiation is the dominant source for sunlit pixels, and diffuse solar radiation for shadowed pixels. A mixed pixel is then resolved by summing up the spectral contribution of all possible light paths.

S3AM embeds a DSM generated by multiview stereo images. The sky view factor F, which is essential to estimate diffuse solar illumination in S3AM, can be conveniently computed using the DSM, reducing the model complexity. Moreover, we take into account the spatial relationship of abundances through weighted TV constraints, derived by spectral information from the hyperspectral imagery, height information from the DSM, and shadow information. The obtained optimization problem is biconvex and is split into two convex problems, separately solved by the ADMM approach iteratively.

The proposed method has been extensively evaluated using two datasets, both quantitatively and qualitatively. Experiments demonstrate that the proposed method significantly reduces the noise level of abundance maps and improves the abundance estimation. Moreover, an ablation study is performed in which the proposed weighted TV constraint is compared to different variations of the spatial TV regularization, or only considering the spectral information.

Several open questions remain. First of all, in order to quantitatively validate shadow-aware spectral unmixing methods in real scenarios, there is a lack of real datasets with shadows for which ground-truth abundances are known. Furthermore, although DSM derived by multiview stereo imagery offers shadow-insensitive height information, the performance of S3AM can be degraded by noise and inaccurate values, especially on boundary regions with large height variations, causing imprecise TV weights and abundance estimation. In particular, urban areas represent the ideal application for this kind of analysis, as shadows are relevant and present across the image. However, DSMs can lead to inconsistencies due to occlusions, especially if the elevation model is derived by stereo images rather than multiview datasets. Thus, future work should address spectral unmixing methods robust to inconsistencies and missing data in the DSM.

APPENDIX
SOLUTION OF ALGORITHM 1

This section presents in detail the updating equations for primal and dual variables in Algorithm 1.

A. Update $\boldsymbol{A}$ and $\boldsymbol{G}$

Given $\boldsymbol{Q}$, $\boldsymbol{K}$, and $\boldsymbol{H}$, the optimization problem in (18) can be rewritten as follows:

$$\min_{\boldsymbol{A},\boldsymbol{G}} \frac{1}{2}\sum_{j=1}^{N}\|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j-\boldsymbol{x}_j\|_F^2+\lambda\|\boldsymbol{G}_2\|_{1,1}+\ell_{\mathcal{C}}(\boldsymbol{G}_3)+\ell_{\mathcal{S}}(\boldsymbol{G}_4)$$
$$\text{s.t.}\begin{cases}\boldsymbol{G}_1=\boldsymbol{A}\\ \boldsymbol{G}_2=\boldsymbol{G}_1\boldsymbol{W}_1\\ \boldsymbol{G}_3=\boldsymbol{A}\\ \boldsymbol{G}_4=\boldsymbol{A}\end{cases} \tag{22}$$

whose augmented Lagrangian is

$$\begin{aligned}\min_{\boldsymbol{A},\boldsymbol{G}} &\frac{1}{2}\sum_{j=1}^{N}\|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j-\boldsymbol{x}_j\|_F^2+\lambda\|\boldsymbol{G}_2\|_{1,1}+\ell_{\mathcal{C}}(\boldsymbol{G}_3)+\ell_{\mathcal{S}}(\boldsymbol{G}_4)\\ &+\frac{\mu}{2}\|\boldsymbol{A}-\boldsymbol{G}_1-\boldsymbol{U}_1\|_F^2+\frac{\mu}{2}\|\boldsymbol{G}_1\boldsymbol{W}_1-\boldsymbol{G}_2-\boldsymbol{U}_2\|_F^2\\ &+\frac{\mu}{2}\|\boldsymbol{A}-\boldsymbol{G}_3-\boldsymbol{U}_3\|_F^2+\frac{\mu}{2}\|\boldsymbol{A}-\boldsymbol{G}_4-\boldsymbol{U}_4\|_F^2.\end{aligned} \tag{23}$$

Thus, we can derive the optimizations with respect to $\boldsymbol{a}_j^{(t+1)}$

$$\boldsymbol{a}_j^{(t+1)}=\left(\widetilde{\boldsymbol{E}}_j^T\widetilde{\boldsymbol{E}}_j+3\mu\boldsymbol{I}\right)^{-1}\left(\widetilde{\boldsymbol{E}}_j^T\boldsymbol{x}_j+\mu\left(\boldsymbol{J}_{1_j}^{(t)}+\boldsymbol{J}_{3_j}^{(t)}+\boldsymbol{J}_{4_j}^{(t)}\right)\right) \tag{24}$$

where $\boldsymbol{J}_{1_j}^{(t)}=\boldsymbol{G}_{1_j}^{(t)}+\boldsymbol{U}_{1_j}^{(t)}$, $\boldsymbol{J}_{3_j}^{(t)}=\boldsymbol{G}_{3_j}^{(t)}+\boldsymbol{U}_{3_j}^{(t)}$, and $\boldsymbol{J}_{4_j}^{(t)}=\boldsymbol{G}_{4_j}^{(t)}+\boldsymbol{U}_{4_j}^{(t)}$.

Next, the optimizations with respect to $\boldsymbol{G}_1^{(t+1)}$, $\boldsymbol{G}_2^{(t+1)}$, $\boldsymbol{G}_3^{(t+1)}$, and $\boldsymbol{G}_4^{(t+1)}$ are written as follows:

$$\boldsymbol{G}_1^{(t+1)}=\left[\boldsymbol{A}^{(t+1)}-\boldsymbol{U}_1^{(t)}+\left(\boldsymbol{G}_2^{t}+\boldsymbol{U}_2^{(t)}\right)\boldsymbol{W}_1^T\right]\left[\boldsymbol{I}+\boldsymbol{W}_1\boldsymbol{W}_1^T\right]^{-1} \tag{25}$$

$$\boldsymbol{G}_2^{(t+1)}=\text{soft}\left(\boldsymbol{G}_1^{(t)}\boldsymbol{W}_1-\boldsymbol{U}_2^{(t)},\frac{\lambda}{\mu}\right) \tag{26}$$

$$\boldsymbol{G}_3^{(t+1)}=\max\left(\boldsymbol{A}^{(t+1)}-\boldsymbol{U}_3^{(t)},0\right) \tag{27}$$

$$\boldsymbol{G}_4^{(t+1)}=\left(\boldsymbol{A}^{(t+1)}-\boldsymbol{U}_4^{(t)}\right)+\frac{1}{p}\left[\boldsymbol{1}_N^T-\boldsymbol{1}_p^T\left(\boldsymbol{A}^{(t+1)}-\boldsymbol{U}_4^{(t)}\right)\right]\otimes\boldsymbol{1}_p. \tag{28}$$

B. Update $\boldsymbol{Q}$, $\boldsymbol{K}$, and $\boldsymbol{H}$

Given $\boldsymbol{A}$ and $\boldsymbol{G}$, the optimization problem in (18) can be rewritten as follows:

$$\min_{\boldsymbol{Q},\boldsymbol{K},\boldsymbol{H}} \frac{1}{2}\sum_{j=1}^{N}\|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j-\boldsymbol{x}_j\|_F^2+\ell_{\mathcal{M}}(\boldsymbol{H}_1)+\lambda\|\boldsymbol{H}_3\|_{1,1}+\ell_{\mathcal{M}}(\boldsymbol{H}_4)$$
$$\text{s.t.}\begin{cases}\boldsymbol{H}_1=\boldsymbol{Q}\\ \boldsymbol{H}_2=\boldsymbol{K}\\ \boldsymbol{H}_3=\boldsymbol{H}_2\boldsymbol{W}_2\\ \boldsymbol{H}_4=\boldsymbol{K}\end{cases} \tag{29}$$

whose augmented Lagrangian is

$$\begin{aligned}\min_{\boldsymbol{Q},\boldsymbol{K},\boldsymbol{H}} &\frac{1}{2}\sum_{j=1}^{N}\|\widetilde{\boldsymbol{E}}_j\boldsymbol{a}_j-\boldsymbol{x}_j\|_F^2+\ell_{\mathcal{M}}(\boldsymbol{H}_1)+\lambda\|\boldsymbol{H}_3\|_{1,1}\\ &+\ell_{\mathcal{M}}(\boldsymbol{H}_4)+\frac{\mu}{2}\|\boldsymbol{Q}-\boldsymbol{H}_1-\boldsymbol{U}_5\|_F^2\\ &+\frac{\mu}{2}\|\boldsymbol{K}_1-\boldsymbol{H}_2-\boldsymbol{U}_6\|_F^2\\ &+\frac{\mu}{2}\|\boldsymbol{H}_2\boldsymbol{W}_2-\boldsymbol{H}_3-\boldsymbol{U}_7\|_F^2\\ &+\frac{\mu}{2}\|\boldsymbol{K}-\boldsymbol{H}_4-\boldsymbol{U}_8\|_F^2.\end{aligned}$$

The optimizations with respect to $\boldsymbol{Q}$ and $\boldsymbol{K}$ are solved pixelwise. For pixel j, we update $Q_j^{(t+1)}$ and $K_j^{(t+1)}$ using

$$Q_j^{(t+1)}=\frac{C_3C_4-C_2C_5}{C_1C_4-C_2^2} \tag{30}$$

$$K_j^{(t+1)}=\frac{C_2C_3-C_1C_5}{C_2^2-C_1C_4} \tag{31}$$

where

$$\begin{cases} C_1 = \sum_{i=1}^{B} (\widetilde{f}_{j,i} - 1)^2 y_{j,i}^2 + \mu \\ C_2 = \sum_{i=1}^{B} (\widetilde{f}_{j,i} - 1)\chi_{j,i} y_{j,i}^2 \\ C_3 = \sum_{i=1}^{B} (\widetilde{f}_{j,i} - 1) y_{j,i}(x_{j,i} - y_{j,i}) + \mu \boldsymbol{J}_{5_j}^{(t)} \\ C_4 = \sum_{i=1}^{B} \chi_{j,i}^2 y_{j,i}^2 + 2\mu \\ C_5 = \sum_{i=1}^{B} \chi_{j,i} y_{j,i}(x_{j,i} - y_{j,i}) + \mu \left(\boldsymbol{J}_{6_j}^{(t)} + \boldsymbol{J}_{8_j}^{(t)} \right) \end{cases} \tag{32}$$

with $\boldsymbol{y}_j = \boldsymbol{E}_j \boldsymbol{a}_j^{(t+1)}$, $\boldsymbol{J}_{5_j}^{(t)} = \boldsymbol{H}_{1_j}^{(t)} + \boldsymbol{U}_{5_j}^{(t)}$, $\boldsymbol{J}_{6_j}^{(t)} = \boldsymbol{H}_{2_j}^{(t)} + \boldsymbol{U}_{6_j}^{(t)}$, and $\boldsymbol{J}_{8_j}^{(t)} = \boldsymbol{H}_{4_j}^{(t)} + \boldsymbol{U}_{8_j}^{(t)}$.

Next, we optimize the objective function with respect to $\boldsymbol{H}_1^{(t+1)}$, $\boldsymbol{H}_2^{(t+1)}$, $\boldsymbol{H}_3^{(t+1)}$, and $\boldsymbol{H}_4^{(t+1)}$ using the following equations:

$$\boldsymbol{H}_1^{(t+1)} = \min\left(\max\left(\boldsymbol{Q}^{(t+1)} - \boldsymbol{U}_5^{(t)}, 0\right), 1\right) \tag{33}$$

$$\boldsymbol{H}_2^{(t+1)} = \left[\boldsymbol{K}^{(t+1)} - \boldsymbol{U}_6^{(t)} + (\boldsymbol{H}_3^t + \boldsymbol{U}_7^{(t)})\boldsymbol{W}_2^T\right]\left[\boldsymbol{I} + \boldsymbol{W}_2\boldsymbol{W}_2^T\right]^{-1} \tag{34}$$

$$\boldsymbol{H}_3^{(t+1)} = \text{soft}\left(\boldsymbol{H}_2^{(t)}\boldsymbol{W}_2 - \boldsymbol{U}_7^{(t)}, \frac{\lambda}{\mu}\right) \tag{35}$$

$$\boldsymbol{H}_4^{(t+1)} = \min\left(\max\left(\boldsymbol{K}^{(t+1)} - \boldsymbol{U}_8^{(t)}, 0\right), 1\right). \tag{36}$$

C. Update $\boldsymbol{U}^{(t+1)}$

$$\begin{cases} \boldsymbol{U}_1^{(t+1)} = \boldsymbol{U}_1^{(t)} - \boldsymbol{A}^{(t+1)} + \boldsymbol{G}_1^{(t+1)} \\ \boldsymbol{U}_2^{(t+1)} = \boldsymbol{U}_2^{(t)} - \boldsymbol{G}_1^{(t+1)} W_1 + \boldsymbol{G}_2^{(t+1)} \\ \boldsymbol{U}_3^{(t+1)} = \boldsymbol{U}_3^{(t)} - \boldsymbol{A}^{(t+1)} + \boldsymbol{G}_3^{(t+1)} \\ \boldsymbol{U}_4^{(t+1)} = \boldsymbol{U}_4^{(t)} - \boldsymbol{A}^{(t+1)} + \boldsymbol{G}_4^{(t+1)} \\ \boldsymbol{U}_5^{(t+1)} = \boldsymbol{U}_5^{(t)} - \boldsymbol{Q}^{(t+1)} + \boldsymbol{H}_1^{(t+1)} \\ \boldsymbol{U}_6^{(t+1)} = \boldsymbol{U}_6^{(t)} - \boldsymbol{K}^{(t+1)} + \boldsymbol{H}_2^{(t+1)} \\ \boldsymbol{U}_7^{(t+1)} = \boldsymbol{U}_7^{(t)} - \boldsymbol{H}_2^{(t+1)}\boldsymbol{W}_2 + \boldsymbol{H}_3^{(t+1)} \\ \boldsymbol{U}_8^{(t+1)} = \boldsymbol{U}_8^{(t)} - \boldsymbol{K}^{(t+1)} + \boldsymbol{H}_4^{(t+1)}. \end{cases} \tag{37}$$

ACKNOWLEDGMENT

The authors would like to thank Dr. Pablo d'Angelo and Dr. Franz Kurz for providing the Digital Surface Models, the Optical Airborne Remote Sensing and Calibration Homebase (OpAIRS) Team with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Weßling, Germany, for providing and preprocessing the HySpex dataset; Dr. Rudolf Richter for assistance with the atmospheric correction of hyperspectral images; and Dr. Bin Yang for sharing the shadow-compensated bilinear mixing model (SCBMM) code. They greatly appreciate anonymous reviewers for their insightful comments and suggestions that greatly improved the quality of this article.

REFERENCES

[1] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1844–1868, Jun. 2014.

[2] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, Jan. 2014.

[3] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: A review," *IEEE Access*, vol. 6, pp. 14118–14129, 2018.

[4] J. M. Bioucas-Dias et al., "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[5] R. Heylen, V. Andrejchenko, Z. Zahiri, M. Parente, and P. Scheunders, "Nonlinear hyperspectral unmixing with graphical models," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4844–4856, Jul. 2019.

[6] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[7] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," *Proc. SPIE*, vol. 7477, Sep. 2009, Art. no. 74770I.

[8] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *Int. J. Remote Sens.*, vol. 30, no. 11, pp. 2951–2962, Jun. 2009.

[9] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[10] R. Heylen and P. Scheunders, "A multilinear mixing model for nonlinear spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 240–251, Jan. 2016.

[11] A. Marinoni and P. Gamba, "A novel approach for efficient p-linear hyperspectral unmixing," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1156–1168, Sep. 2015.

[12] J. R. Schott, *Remote Sensing: The Image Chain Approach*. New York, NY, USA: Oxford Univ. Press, 2007.

[13] S. M. Adler-Golden, M. W. Matthew, G. P. Anderson, G. W. Felde, and J. A. Gardner, "An algorithm for de-shadowing spectral imagery," *Proc. SPIE*, vol. 4816, pp. 203–210, Nov. 2002.

[14] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 650–663, Mar. 2004.

[15] M. Zhao, J. Chen, and S. Rahardja, "Hyperspectral shadow removal via nonlinear unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 5, pp. 881–885, May 2021.

[16] T. Uezato, N. Yokoya, and W. He, "Illumination invariant hyperspectral image unmixing based on a digital surface model," *IEEE Trans. Image Process.*, vol. 29, pp. 3652–3664, 2020.

[17] G. Zhang, D. Cerra, and R. Müller, "Shadow detection and restoration for hyperspectral images based on nonlinear spectral unmixing," *Remote Sens.*, vol. 12, no. 23, p. 3985, Dec. 2020.

[18] B. Yang, "Supervised nonlinear hyperspectral unmixing with automatic shadow compensation using multiswarm particle swarm optimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5529618.

[19] G. Zhang, P. Scheunders, D. Cerra, and R. Müller, "Shadow-aware nonlinear spectral unmixing for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 5514–5533, 2022.

[20] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization," *Swarm Intell.*, vol. 1, no. 1, pp. 33–57, Jun. 2007.

[21] C. Shi and L. Wang, "Incorporating spatial information in spectral unmixing: A review," *Remote Sens. Environ.*, vol. 149, pp. 70–87, Jun. 2014.

[22] L. Wang, C. Shi, C. Diao, W. Ji, and D. Yin, "A survey of methods incorporating spatial information in image classification and spectral unmixing," *Int. J. Remote Sens.*, vol. 37, no. 16, pp. 3870–3910, Aug. 2016.

[23] S. Lefkimmiatis and S. Osher, "Nonlocal structure tensor functionals for image regularization," *IEEE Trans. Comput. Imag.*, vol. 1, no. 1, pp. 16–29, Mar. 2015.

[24] J. Yao, D. Meng, Q. Zhao, W. Cao, and Z. Xu, "Nonconvex-sparsity and nonlocal-smoothness-based blind hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2991–3006, Jun. 2019.

[25] F. van der Meer, "Iterative spectral unmixing (ISU)," *Int. J. Remote Sens.*, vol. 20, no. 17, pp. 3431–3436, Jan. 1999.

[26] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Enhancing hyperspectral image unmixing with spatial correlations," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4239–4247, Nov. 2011.

[27] J. Yao, D. Hong, L. Xu, D. Meng, J. Chanussot, and Z. Xu, "Sparsity-enhanced convolutional decomposition: A novel tensor-based paradigm for blind hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5505014.

[28] D. Hong et al., "Endmember-guided unmixing network (EGU-Net): A general deep learning framework for self-supervised hyperspectral unmixing," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 11, pp. 6518–6531, Nov. 2022.

[29] X. Song, X. Jiang, and X. Rui, "Spectral unmixing using linear unmixing under spatial autocorrelation constraints," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 975–978.

[30] A. Castrodad, Z. Xing, J. B. Greer, E. Bosch, L. Carin, and G. Sapiro, "Learning discriminative sparse representations for modeling, source separation, and mapping of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4263–4281, Nov. 2011.

[31] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 525–538, Jan. 2016.

[32] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.

[33] T. Uezato, M. Fauvel, and N. Dobigeon, "Hyperspectral image unmixing with LiDAR data-aided spatial regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 4098–4108, Jul. 2018.

[34] J. Liu, J. Zhang, Y. Gao, C. Zhang, and Z. Li, "Enhancing spectral unmixing by local neighborhood weights," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 5, pp. 1545–1552, Oct. 2012.

[35] Y. Zhong, R. Feng, and L. Zhang, "Non-local sparse unmixing for hyperspectral remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1889–1909, Jun. 2014.

[36] 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.

[37] B. Yang and B. Wang, "Band-wise nonlinear unmixing for hyperspectral imagery using an extended multilinear mixing model," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6747–6762, Nov. 2018.

[38] J. Yang, Y. He, and J. Caspersen, "Fully constrained linear spectral unmixing based global shadow compensation for high resolution satellite imagery of urban areas," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 38, pp. 88–98, Jun. 2015.

[39] F. A. Kruse et al., "The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data," *Remote Sens. Environ.*, vol. 44, nos. 2–3, pp. 145–163, May/Jun. 1993. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0034425793900013N

[40] F. Yamazaki, W. Liu, and M. Takasaki, "Characteristics of shadow and removal of its effects for remote sensing imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 4, Jul. 2009, pp. IV-426–IV-429.

[41] D. Cerra et al., "DLR HySU—A benchmark dataset for spectral unmixing," *Remote Sens.*, vol. 13, no. 13, p. 2559, Jun. 2021.

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

[43] O. Conrad et al., "System for automated geoscientific analyses (SAGA) v. 2.1.4," *Geosci. Model Develop.*, vol. 8, no. 7, pp. 1991–2007, Jul. 2015.

[44] C. H. Köhler, "Airborne imaging spectrometer HySpex," *J. Large-Scale Res. Facilities JLSRF*, vol. 2, no. A93, pp. 1–6, Nov. 2016.

[45] R. Richter and D. Schläpfer, *Atmospheric/Topographic Correction for Airborne Imagery*, document 565-02, ATCOR-4 User Guide, 2011.

[46] B. Somers, M. Zortea, A. Plaza, and G. P. Asner, "Automated extraction of image-based endmember bundles for improved spectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 396–408, Apr. 2012.

[47] P. d'Angelo and F. Kurz, "Aircraft based real time bundle adjustment and digital surface model generation," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 42, pp. 1643–1647, Jun. 2019.

Guichen Zhang received the B.Sc. degree in remote sensing science and technology from Wuhan University, Wuhan, China, in 2014, the M.Sc. degree in earth-oriented space science and technology from the Technical University of Munich, Munich, Germany, in 2018, and the M.Sc. degree in photogrammetry and remote sensing from Wuhan University in 2018. She is currently pursuing the Ph.D. degree with Osnabrück University, Osnabrück, Germany.

In 2021, she was a Visiting Scientist with the Physics Department, University of Antwerp, Antwerp, Belgium. Since 2018, she has been a Research Associate with the German Aerospace Center, Weßling, Germany. Her research interests include image processing, hyperspectral image analysis, data fusion, and machine learning in the application of remote sensing.

Paul Scheunders (Senior Member, IEEE) received the B.S. and Ph.D. degrees in physics from the University of Antwerp, Antwerp, Belgium, in 1983 and 1990, respectively, with a focus in statistical mechanics.

In 1991, he became a Research Associate with the Vision Lab, Department of Physics, University of Antwerp, where he is currently a Full Professor. He has published over 200 papers in international journals and proceedings in the field of image processing, pattern recognition, and remote sensing. His research interests include remote sensing and hyperspectral image processing.

Dr. Scheunders is a Senior Member of the IEEE Geoscience and Remote Sensing Society (GRSS). He has served as a program committee member in numerous international conferences. He is an Associate Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Daniele Cerra (Member, IEEE) received the Ph.D. degree (Hons.) in image and signal processing from Télécom Paristech University, Paris, France, in 2010, and two M.Sc. degrees in computer engineering and geographic information system (GIS) from Roma Tre University, Rome, Italy, in 2005, and Salamanca University, University, Spain, in 2006.

He has been with the Department of Photogrammetry and Image Analysis, German Aerospace Center (DLR), Oberpfaffenhofen, Weßling, Germany, since 2007. He has been teaching digital image processing and hyperspectral remote sensing since 2012 as a Visiting Lecturer or a Professor in several universities (Spain, Germany, Italy, and Iran) and since 2014 as a Trainer at the European Space Agency (ESA), Noordwijk, Netherlands. He coauthored the best paper in the IEEE Whispers 2019 Conference. He authored or coauthored more than 100 publications in peer-reviewed journals, books, or international conferences. His research interests include hyperspectral imaging, data fusion, and algorithmic information theory.

Dr. Cerra won a total of five prizes (three as first author) in the IEEE Geoscience and Remote Sensing Society (GRSS) Data Fusion Contests in the years 2013–2020 and the DLR Science Slam Contest in 2013.