Snapshot Multispectral Image Completion and Unmixing with Total Variation Regularization on Abundance Maps

Keisuke Ozawa*, Shinichi Sumiyoshi*, Yuki Tachioka*
* DENSO IT Laboratory, Tokyo, Japan
E-mail: ozawa.keisuke@core.d-itlab.co.jp

***Abstract*—Unmixing is an important application of spectral imaging, and snapshot sensors could enrich its applicability. However, their spatio-spectral tradeoff decreases spatial resolution as the number of bands increases. While basis spectra can be estimated even on the downsampled multispectral image, it is difficult to retain high-resolution abundance maps. In this paper, we propose a high spatial resolution unmixing method from a single snapshot multispectral image. The proposed method simultaneously completes a snapshot data to restore the full sensor size multispectral image. In a simulation, we show a resolution-enhanced unmixing and better completion accuracy compared with state-of-the-art tensor completion methods. We also demonstrate against real data the best quality for completion and unmixing in the full sensor size.**

I. Introduction

Multispectral unmixing is a source separation technique widely studied in remote sensing [1]. It decomposes a spectral image into basis spectra or endmembers, often representing composite materials, and their abundance maps. A basis spectrum is the reflectance of some pure material and the abundance map is its spatial distribution. The applicability of unmixing will be enriched by snapshot imaging with mosaic filter arrays [2], [3] that can capture dynamic scenes with multispectral resolution [4], [5], [6]. However, the raw signal is equivalent to a sparsely sampled tensor with its sensor size and the number of bands ($\mathcal{M}$ in Fig. 1). Mosaic rearrangement reconstructs a low resolution multispectral image but spatially down-samples and degrades fine textures. This spatio-spectral tradeoff prevents the capture of a high-resolution multispectral image and thus abundance maps from a snapshot.

An accurate enough completion method, if it exists, would resolve our problem, followed by an additional unmixing process. However, existing completion methods have some limitations. Demosaicing methods [7], [8], [9], [10], [11] assume a small number of channels or require specific spectral profiles and arrangement to satisfy inter-channel correlations. Neither linear interpolation nor inter-channel correlation can be applied to our study, because we use built-in filters with distinct responses. Low rank assumption is known as useful for tensor recovery, especially because it is often assumed that the rich spectral information in spectral imaging has low-rank structures. Many types of tensor factorizations have been proposed. Tensor unfolding into matrices [12], with the total variation regularization [13], might ignore correlation among spatial dimensions. Tensor ring [14] entangles dimensions of different meanings and might be difficult to fit with the data structure of multispectral images. Tensor tubal rank [15] considers inter-axis correlation, and t(transformed)-TNNs as its convex relaxation have achieved state-of-the-art results [16], [17], [18], [19], [20], [21], [22]. However, we consider that the low-rank approximation loses some frequency information in reconstructed spectra. As a result, existing completion methods are insufficient for the purpose of high-resolution restoration from snapshot multispectral images and further unmixing.

To tackle this problem, we propose a simultaneous completion and unmixing method to obtain a high-resolution multispectral image and abundance maps from a snapshot. Fig. 1 illustrates our idea. We facilitate the mosaic-rearranged multispectral image to estimate basis spectra. We assume that signal contamination on spectra introduced by pixels' misalignment from the mosaic rearrangement is moderate enough for endmeber extraction, e.g. by using the vertex component analysis (VCA) [23][1]. Note that we do not consider whether they are the (pure) emdmembers or not; we just consider basis spectra for the linear mixing model. Then, our idea is to complete missing entries by linear mixing of the basis spectra while assuring smoothness of the abundance maps. The basis spectra, the observed tensor entries, and the nonnegativity constraint determine the solution spaces of the tensor and the abundance maps, where we prefer the most smoothly varying abundance maps. We are inspired by the previous work [25], [26], [27], [28] to regularize smoothness of abundance maps. We use the anisotropic total variation (TV) [29] as in [25], [26], but it can be replaced by the isotropic TV [30] or the other smoothness regularizations. We expect that optimizing over the solution spaces is useful to obtain high-resolution abundance maps and better complete snapshot data, compared to existing tensor completion methods followed by unmixing.

In the remainder of this paper, we summarize the notations, formulate our idea as an optimization problem, and then derive update rules by the alternating direction method of multipliers (ADMM) [31]. We conduct a simulation experiment to evaluate the completion accuracy and show resolution-enhanced abundance maps. We also demonstrate quality enhancement on full sensor size completion and unmixing on real data.

[1]Robust PCA [24] could also be used to reduce outliers in advance.

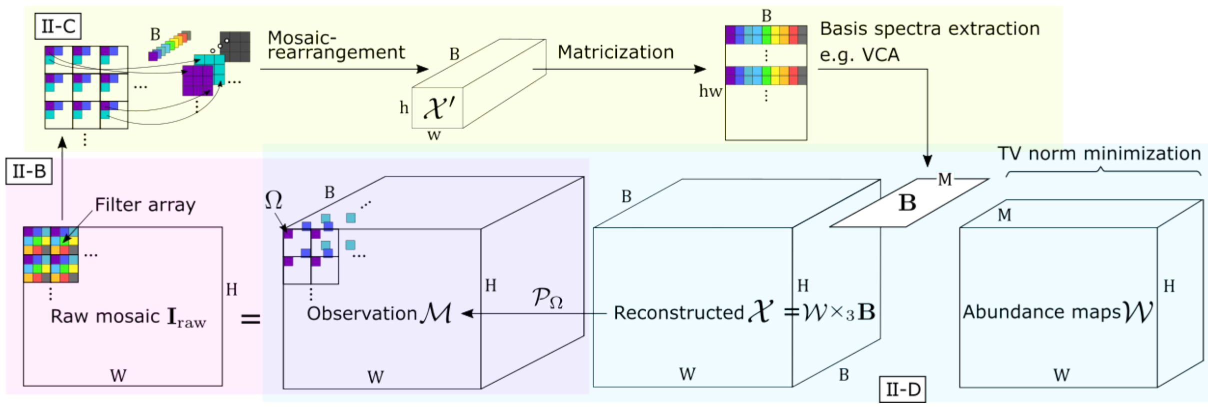

Fig. 1. Schematic of the proposed method.

II. PROPOSED METHOD

A. Notations

A multispectral image is a three-way tensor, denoted like $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ in calligraphy. Its k-th frontal slice is a matrix $\mathbf{A}^{(k)} = \mathcal{A}(:,:,k) \in \mathbb{R}^{n_1 \times n_2}$ in bold letters. The symbol : means that all the elements appear in that dimension. $\|\mathcal{A}\|_F^2 = \sum_{i,j,k} |\mathcal{A}(i,j,k)|^2$ is the square of the Frobenius norm of the tensor $\mathcal{A}$. The (i,j)-th tube of $\mathcal{A}$ is a vector $\mathbf{a}^{(ij)} \in \mathbf{R}^{n_3}$, and the tubal transformation [15] by $\mathbf{P} \in \mathbf{R}^{n_3' \times n_3}$ is $\mathcal{A} \times_3 \mathbf{P} \in \mathbb{R}^{n_1 \times n_2 \times n_3'}$, for which $\mathcal{A} \times_3 \mathbf{P}(i,j,:) = \mathbf{P}\mathbf{a}^{(ij)} \in \mathbf{R}^{n_3'}$. The anisotropic TV norm [29] of $\mathbf{A}$ is $\|\mathbf{A}\|_{\mathrm{TV}} := \sum_{i,j} |\partial_x \mathbf{A}_{ij}| + |\partial_y \mathbf{A}_{ij}|$, where $\nabla = (\partial_x ; \partial_y)$ is the differential operator and ; vertically concatenates the operators. $\mathcal{A} \succeq 0$ has nonnegative elements. The projection of $\mathcal{A}$ onto $\Omega \subset \mathrm{dom}(\mathcal{A})$ has $[\mathcal{P}_\Omega(\mathcal{A})](i,j,k) = \mathcal{A}(i,j,k)$ if $(i,j,k) \in \Omega$ and 0 otherwise. $\mathbb{R}_{\geq 0}$ is the set of nonnegative real numbers. The indicator function on Ω gives $\iota_\Omega(\mathcal{A};\mathcal{M}) = 0$ if $\mathcal{P}_\Omega(\mathcal{A}) = \mathcal{P}_\Omega(\mathcal{M})$ for $\mathcal{M}$ that observes the elements in Ω, and ∞ otherwise. The soft-thresholding operator [32] acts element-wise as $\mathcal{S}_\lambda(x) = \mathrm{sgn}(x)\max(|x| - \lambda, 0)$.

B. Preparing observed tensor

The sensor are deposited with repeated $b \times b$ filters of different transmission spectra on its surface. Every pixel has one filter on it. A snapshot by the sensor is a mosaic image $\mathbf{I}_{\mathrm{raw}} \in \mathbb{R}^{H \times W}$ with the repeated $b \times b$ spectral bands (Fig. 1 II-B). We convert this raw image into an observed tensor $\mathcal{M} \in \mathbb{R}^{H \times W \times B}$ as $\mathcal{M}(i,j,k(i(\mathrm{mod}\ b), j(\mathrm{mod}\ b))) = \mathbf{I}_{\mathrm{raw}}(i,j)$, where k is the band index of $b \times b (= B)$ filters. $\mathcal{M}$ has observed values sparsely only at the elements of Ω, where the sampling rate is $1/B$ and one band observation at a pixel.

C. Mosaic-rearrangement and basis spectra extraction

We can roughly reconstruct a spectrum by vectorizing the $b \times b$ pixels into a spectrum of B bands (Fig. 1 II-C). As described in Section I, such a reconstructed spectrum deviates from the true spectrum due to pixels' misalignment at material boundaries and shading change. However, spectra are reasonably recovered when $b \times b$ pixels observe an equally illuminated surface region composed of the same material. We assume scenes wherein changes of local textures and shading are not so drastic within most of the $b \times b$ pixel regions over the image. We expect that the basis spectra are obtained with near correct spectra by reducing such moderate spectral deviations.

Under these assumptions, we mosaic-rearrange the observed tensor $\mathcal{M}$ (or equivalently the raw mosaic image) into a spatially down-sampled tensor $\mathcal{X}' \in \mathbb{R}^{h \times w \times B}$, where $h = H/b$ and $w = W/b$, respectively (Fig. 1 II-C). It follows $\mathcal{X}'(i',j',k) = I((i'-1)b + k_x^{-1}, (j'-1)b + k_y^{-1})$, where the abbreviated symbol $k_{\{x,y\}}^{-1}$ denotes $\{x,y\}$-position that corresponds to the band index k. Then, we extract basis spectra on the mosaic-rearranged, low-resolution multispectral image $\mathcal{X}'$. In this study, we use VCA algorithm [23] (The details are in the experimental section). Then, the power of the extracted spectra are normalized to one. We denote the basis spectra as $\mathbf{B} \in \mathbb{R}^{B \times M}$, each column of which is an basis spectrum.

D. Problem Formulation

We will optimize the full sensor size multispectral image $\mathcal{X} \in \mathbb{R}^{H \times W \times B}$ and the abundance maps $\mathcal{W} \in \mathbb{R}^{H \times W \times M}$ simultaneously, where H and W are the height and width of the sensor, B is the number of bands, M is the number of basis spectra, and each frontal-slice of $\mathcal{W}$ is an abundance map (Fig. 1 II-D). We minimize their TV norm [29] as follows:

$$\{\mathcal{X}^*, \mathcal{W}^*\} = \arg\min_{\mathcal{X},\mathcal{W}} \sum_{m=1}^{M} \|\mathbf{W}^{(m)}\|_{\mathrm{TV}}$$
$$\text{s.t. } \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{M}),\ \mathcal{X} = \mathcal{W} \times_3 \mathbf{B},\ \mathcal{W} \succeq 0. \quad (1)$$

The first constraint requires the recovered data and that the sensor values are consistent at the observed elements. The second constraint implies that we follow the linear mixing model for multispectral images: $\mathcal{X}$ and $\mathcal{W}$ should be consistent to each other to satisfy this constratint. The third constraint is the nonnegativity constraint of the abundance maps. Hereafter, we denote $\tilde{\mathcal{X}} := \mathcal{W} \times_3 \mathbf{B}$ for brevity. Note that we require no tuning parameter and will obtain the smoothest abundance maps among possible solutions under the constraints.

E. Optimization

We solve the convex optimization problem (1) with the ADMM framework in a similar manner to [31]. The augmented Lagrangian is

$$
\begin{aligned}
&\mathcal{L}(\mathcal{X},\mathcal{W};\mathcal{U},\mathcal{V}) = \iota_{\Omega}(\mathcal{X};\mathcal{M}) + \iota_{\mathbb{R}^{H\times W\times M}_{\geq 0}}(\mathcal{W}) \\
&+ \sum_{m=1}^{M} \|\mathbf{V}_x^{(m)}\|_1 + \sum_{m=1}^{M} \|\mathbf{V}_y^{(m)}\|_1 + \langle \mathcal{X} - \tilde{\mathcal{X}}, \Phi_1 \rangle \\
&+ \frac{\rho_1}{2}\|\mathcal{X} - \tilde{\mathcal{X}}\|_F^2 + \langle \mathcal{W} - \mathcal{U}, \Phi_2 \rangle + \frac{\rho_2}{2}\|\mathcal{W} - \mathcal{U}\|_F^2 \\
&+ \sum_{m=1}^{M} \left(\langle \nabla \mathbf{U}^{(m)} - \mathbf{V}^{(m)}, \mathbf{\Phi_3}^{(m)} \rangle + \frac{\rho_3}{2}\|\nabla \mathbf{U}^{(m)} - \mathbf{V}^{(m)}\|_F^2 \right),
\end{aligned}
\tag{2}
$$

where $\mathcal{U}$ and $\mathcal{V}$ are the auxiliary variables, $\Phi_i (i = 1, 2, 3)$ are the Lagrange multipliers, and $\mathbf{\Phi}_3^{(m)}$ is the (m)-th frontal slice matrix of Φ_3. Note that $\nabla \mathbf{U}^{(m)} := (\partial_x; \partial_y)\mathbf{U}^{(m)}$, where we assume the periodic boundary condition, and $\mathbf{V}^{(m)} = (\mathbf{V}_x^{(m)}; \mathbf{V}_y^{(m)}) \in \mathbb{R}^{2H\times W}$. We hereafter add the superscript (t) to parameters updated in the t-th iteration. Subproblems at $(t+1)$-th iteration are solved as follows:

1) The update of $\mathcal{X}$ is a square error minimization under the observation, and written in a closed form with a projection on the complement Ω^c of the observed elements Ω:

$$
\begin{aligned}
\mathcal{X}^{(t+1)} &= \arg\min_{\mathcal{X}} \iota_{\Omega}(\mathcal{X};\mathcal{M}) + \frac{\rho_1^{(t)}}{2}\|\mathcal{X} - \left(\tilde{\mathcal{X}}^{(t)} - \frac{\Phi_1^{(t)}}{\rho_1^{(t)}}\right)\|_F^2 \\
&= \mathcal{P}_{\Omega^c}\left(\tilde{\mathcal{X}}^{(t)} - \frac{\Phi_1^{(t)}}{\rho_1^{(t)}}\right) + \mathcal{M}
\end{aligned}
\tag{3}
$$

2) The update of $\mathcal{W}$ is a constrained least square problem:

$$
\begin{aligned}
\mathcal{W}^{(t+1)} &= \arg\min_{\mathcal{W}} \frac{\rho_1^{(t)}}{2}\|\mathcal{X}^{(t+1)} - \left(\tilde{\mathcal{X}}^{(t)} - \frac{\Phi_1^{(t)}}{\rho_1^{(t)}}\right)\|_F^2 \\
&\quad + \iota_{\mathbb{R}^{H\times W\times M}_{\geq 0}}(\mathcal{W}) + \frac{\rho_2^{(t)}}{2}\|\mathcal{W} - \left(\mathcal{U}^{(t)} - \frac{\Phi_2^{(t)}}{\rho_2^{(t)}}\right)\|_F^2 \\
&= \iota_{\mathbb{R}^{H\times W\times M}_{\geq 0}}(\mathcal{W}) + \frac{1}{2}\|\mathcal{W} \times_3 \tilde{\mathbf{B}}^{(t)} - \mathcal{Y}^{(t+1)}\|_F^2,
\end{aligned}
\tag{4}
$$

where $\tilde{\mathbf{B}}^{(t)}$ and $\mathcal{Y}^{(t+1)}$ in the last equation are

$$
\tilde{\mathbf{B}}^{(t)} = \begin{pmatrix} \sqrt{\rho_1^{(t)}}\mathbf{B} \\ \sqrt{\rho_2^{(t)}}\mathbf{I} \end{pmatrix}, \mathcal{Y}^{(t+1)} = \begin{pmatrix} \sqrt{\rho_1^{(t)}}\left(\mathcal{X}^{(t+1)} + \frac{\Phi_1^{(t)}}{\rho_1^{(t)}}\right) \\ \sqrt{\rho_2^{(t)}}\left(\mathcal{U}^{(t)} - \frac{\Phi_2^{(t)}}{\rho_2^{(t)}}\right) \end{pmatrix}
\tag{5}
$$

where $\mathbf{I}$ is the identity matrix. This subproblem with the nonnegative constraint can be solved by an existing implementation [33].

3) The update of $\mathcal{U}$ is

$$
\begin{aligned}
\mathcal{U}^{(t+1)} &= \arg\min_{\mathcal{U}} \frac{\rho_2^{(t)}}{2}\|\mathcal{U} - \left(\mathcal{W}^{(t+1)} + \frac{\Phi_2^{(t)}}{\rho_2^{(t)}}\right)\|_F^2 \\
&+ \frac{\rho_3^{(t)}}{2}\sum_{m=1}^{M}\|\nabla \mathbf{U}^{(m)} - \left(\mathbf{V}^{(m)(t)} - \frac{\mathbf{\Phi}_3^{(m)(t)}}{\rho_3^{(t)}}\right)\|_F^2.
\end{aligned}
\tag{6}
$$

The m-th slice of $\mathcal{U}^{(t+1)}$ satisfies

$$
\begin{aligned}
&(\rho_2^{(t)}\mathbf{1} + \rho_3^{(t)}\nabla^{\mathrm{T}}\nabla)\mathbf{U}^{(m)(t+1)} = \rho_2^{(t)}\mathbf{W}'^{(m)} + \rho_3^{(t)}\nabla^{\mathrm{T}}\mathbf{V}'^{(m)}, \\
&\mathcal{W}' := \mathcal{W}^{(t+1)} + \frac{\Phi_2^{(t)}}{\rho_2^{(t)}},\ \mathcal{V}' := \mathcal{V}^{(t)} - \frac{\Phi_3^{(t)}}{\rho_3^{(t)}},
\end{aligned}
\tag{7}
$$

from which the update can be computed efficiently in the Fourier domain [34].

4) The update of $\mathcal{V}$ is the following L1-norm minimization:

$$
\begin{aligned}
&\mathbf{V}^{(m)(t+1)} \\
&= \arg\min_{\mathbf{V}^{(m)}} \frac{\rho_3^{(t)}}{2}\|\mathbf{V}^{(m)} - \left(\nabla \mathbf{U}^{(m)(t+1)} + \frac{\mathbf{\Phi}_3^{(m)(t)}}{\rho_3^{(t)}}\right)\|_F^2 \\
&+ \|\mathbf{V}_x^{(m)}\|_1 + \|\mathbf{V}_y^{(m)}\|_1 \\
&= \mathcal{S}_{1/\rho_3^{(t)}}\left(\nabla \mathbf{U}^{(m)(t+1)} + \frac{\mathbf{\Phi}_3^{(m)(t)}}{\rho_3^{(t)}}\right)
\end{aligned}
\tag{8}
$$

5) Finally, the updates of the Lagrange multipliers are

$$
\Phi_1^{(t+1)} = \Phi_1^{(t)} + \rho_1^{(t)}(\mathcal{X}^{(t+1)} - \mathcal{W}^{(t+1)} \times_3 \mathbf{B}) \tag{9}
$$

$$
\Phi_2^{(t+1)} = \Phi_2^{(t)} + \rho_2^{(t)}(\mathcal{W}^{(t+1)} - \mathcal{V}^{(t+1)}) \tag{10}
$$

$$
\mathbf{\Phi}_3^{(m)(t+1)} = \mathbf{\Phi}_3^{(m)(t)} + \rho_3^{(t)}(\nabla \mathbf{V}^{(m)(t+1)} - \mathbf{U}^{(m)(t+1)}) \quad \text{for } m = 1, \cdots, M. \tag{11}
$$

$$
\rho_l^{(t+1)} = \min\left\{\gamma\rho_l^{(t)}, \rho_{\max}\right\} \text{ for } l = 1, 2, 3, \tag{12}
$$

where γ accelerates the convergence speed when $\gamma > 1$.

Algorithm 1 Multispectral image completion and unmixing

Input: Observed sparse tensor $\mathcal{M} \in \mathbb{R}^{H\times W\times B}$
Output: Completed multispectral image $\mathcal{X}^* \in \mathbb{R}^{H\times W\times B}$ and abundance maps $\mathcal{W}^* \in \mathbb{R}^{H\times W\times M}$
Initialization : $\mathcal{X} = \mathcal{M}$, $\mathcal{W} = 0$, observed tensor elements Ω, $\rho_i^{(0)}$, $\Phi_i^{(0)} = 0$ $(i = 1, 2, 3)$;
Extract basis spectra with VCA algorithm.
Parameters : γ, tol, max iteration, $\rho_{\max}$
1: **while** below max iteration or max diff $>$ tol **do**
2: Update $\mathcal{X}$ using Eq. (3).
3: Update $\mathcal{W}$ using Eq. (4).
4: Update $\mathcal{U}$ using Eq. (7).
5: Update $\mathcal{V}$ using Eq. (8).
6: Update Lagrange multipliers using Eq. (9)-(12).
7: Continue updates 2-6
8: **end while**

We implemented Algorithm 1 with MATLAB2020b on a 64-bit computer with an Intel Core-i7-8700 3.20GHz CPU.

We take the observed sparse tensor as input with the indices of observed entries, set the basis spectra extracted with the VCA algorithm, and then iteratively update the variables until the maximum absolute change of variables goes lower than the tolerance or the maximum number of iterations is achieved.

III. Experiments

A. Data preparation

Simulation data are prepared for numerical evaluation. We rendered Bunny and Dragon from the Stanford 3D scanning repository [35] in the size of 510×510 pixels and 9 bands. The objects are textured with two or four randomly sampled Macbeth colors. We used a renderer PBRT-v2 [36], which we modified to simulate spectral images. The rendered images are used as the ground truth, and we simulate a snapshot input as of the same tensor structure as the real sensor but with a different filter arrangement and transmittance to simplify simulation. Specifically, we set to 2 nm spectral resolution and 9 bands are chosen starting from 400 nm with 10 nm steps. Interreflection is simulated up to 5 times. Although interreflection violates the assumptions of linear unmixing and should be considered as a weak outlier, we keep them into consideration for a realistic imaging assumption.

Real data are captured with the snapshot camera CMS-C (SILIOS product). It captures a raw sensor image with 12-bit of dynamic range and a size of 1280×1024 pixels. Every 3×3 pixels repeated on the sensor plane (Fig. 1 II-B) are deposited with thin layers of 9 different transmittance ($b = 3$ and $B = 9$ in Section II-B). The raw image is rearranged into a subsampled tensor $\mathcal{M}$ (Fig. 1 II-B) as described in Section II-B. The target object is a dyed fabric, which is of diffuse reflection with less number of materials (or colorants) than the number of bands, so that basic assumptions for linear unmixing holds. There is a need for the flat field correction: first, we take a white reference illuminated by light sources then take the object with the same lighting. (The necessity of flat field correction depends on sensor characteristics.) In this study, we illuminated the object by two halogen lamps placed at both sides of the camera in a dark room.

The following procedures are the same for both simulation and real data, for numerical comparison and qualitative comparison, respectively.

B. Basis spectra extraction

We extracted basis spectra on the rearranged, low-resolution multispectral images (Fig. 1 II-C) $\mathcal{X}'$ with VCA [23] by manually fixing the number of basis spectra M. Note that VCA also removes moderate noise from inputs in its preprocess. We fix the number M that gives the smallest squared error between the rearranged low-resolution image $\mathcal{X}'$ and reconstructed tensor from the basis spectra and low-resolution abundance maps using SUnSAL [33], which unmix a spectral image under the assumptions of linear mixing and nonnegative abundance maps as we also assume. Although VCA finds basis spectra for each trial, we optimize on the extracted spectra for the first trial in this experiment. Note that, except for our motivation to complete tensors, estimated abundance maps on the low-resolution image and our reconstruction are expected to have similar distributions but different resolutions. We will see this expectation experimental results.

Note also that we can use any algorithm to extract basis spectra in our method. Although the appropriate basis spectra extraction depends on the purpose of unmixing, in this study, we resort to VCA widely used for general purposes.

C. Tensor completion and optimization of abundance maps

The input tensor $\mathcal{M}$ is normalized so that elements are within the range of $[0, 1]$. We run Algorithm 1 with fixing parameters: tolerance 10^{-5}; maximum number of iteration 500; the acceleration speed $\gamma = 1.1$; and $\rho_1^{(0)} = \rho_2^{(0)} = \rho_3^{(0)} = 10^{-1}$ for real captured data, $\rho_1^{(0)} = \rho_2^{(0)} = \rho_3^{(0)} = 1$ for simulation data, respectively. These parameters were determined to get a better convergence speed in experiments.

D. Results

We evaluated the completion accuracy of our method against the state-of-the-art completion methods: LRTC-TV [13] that uses low-rankness of unfolded matrices with (isotropic) TV regularization; TRLRF [14] in the framework of tensor ring; DCTNN [18] that uses tensor low-rankness of transformed tensors by discrete cosine transformation (DCT); FTNN [21] that shares the same spirit of DCTNN [18] but introduces framelet transformation to define tensor nuclear norm; and TNTV [22] that augments DCTNN [18] with TV regularization (in addition to allowing Gaussian noise and sparse outliers). Other than DCTNN [18], for which we used our implementation, we used the authors' implementation with Matlab. TNTV [22] provides options whether using DCT or discrete Fourier transformation (DFT), where we only considered DCT since many existing studies have suggested that DCT performs better than DFT for t-TNN families. We used the default tuning parameters recommended by the authors except for FTNN [21], for which we set the level to 3, the maximum level that worked for our inputs. It obtained better results than when the level is set to 2. For TNTV [22], we empirically tuned the weight on TV to increase performance. The other parameters for noise and outliers are default but they are not so sensitive to our results. Note that some of these methods provide ways to automatically determine the tuning parameters from input data.

1) Results for simulation experiment: Table I shows the comparison of completion performance. The reconstruction quality was measured by the relative error (RelE)

$$RelE = \frac{\|\mathcal{X}^\star - \mathcal{X}^0\|_{\mathrm{F}}^2}{\|\mathcal{X}^0\|_{\mathrm{F}}^2}$$

and the peak signal-to-noise ratio (PSNR) averaged over all the band images. Our completion results were the best in terms of RelE. This is considerably because we have facilitated the snapshot structure to extract basis spectra and our formulation could better describe the subspace of multispectral images, which are spanned with the basis spectra and have spatial smoothness. TNTV [22] performed the second, but competed with DCTNN [18] for all the scenes. Although FTNN [14] is reported to perform the better against different tasks, DCTNN [18] was better for our objects. Note that, in this numerical and the next qualitative evaluations, our results do not mean the inferior performance of the existing methods and rather suggest that our formulation has advantages to treat snapshot multispectral images: there are some difficulties for fair enough comparison, e.g., it is often difficult to set the appropriate rank to apply TRLRF. Also, the methods without any smoothness regularization (TRLRF, DCTNN, FTNN) have been studied under random sampling assumptions, while the periodic missing entries of snapshot multispectral imaging might have minimizers that are different from the original images of smooth appearance. LRTC-TV assumes the same smoothness, but again the periodic missing entries might be difficult to complete with this model. In contrast, our method specifically facilitates the data structure to extract basis spectra and we could restrict the solution space.

The quality of our completion results in simulation was better than the other methods (Fig. 3, where the observed band-1 images are almost invisible because the sampling is so sparse). Although TNTV [22] performed the second best in numerical evaluation, DCTNN [18] seems to have produced reasonable results. However, t-TNN only considers the global redundancy of images and thus block artifacts remained (best viewed in zoom in digital format). TNTV [22] additionally considers local smoothness, but t-TNN and the TV regularization are difficult to balance and seems insufficient to characterize multispectral images. This method also oversmoothes over the dark background due to noise-aware data fidelity, while RelE and PSNR got worse if we decrease the smoothness. On the other hand, our method does not require to balance smoothness and simply requires to minimize the TV norms of abundance maps with regressing unobserved tensor values over basis spectra. We consider that this endows a better solution space.

To demonstrate the high-resolution recovery of abundance maps, we also show in the simulation the recovered high-resolution abundance maps in comparison with the low-resolution maps generated by mosaic-rearrangement and SUnSAL algorithm [33] (Fig. 2). The reference color image of the object and two zoomed-up regions of those abundance maps are also shown. Although the detailed textures are corrupted in the low-resolution abundance maps, they can be clearly observed in our reconstructed results.

2) Results for real data experiments: In a real experiment, our method provided a better high-resolution image than the compared methods (Fig. 4). TNTV [22] was the second best as in simulation, and the quality comparison in zoomed-up images demonstrates that TNTV [22] overly smoothed textures and dotted artifacts remained. Our completion results better recovered fine textures. Wrinkles on the background (second and 5,6-th column in the bottom two rows) are well displayed in our result, while it completely disappeared in TNTV [22].

The details of our completion results (Fig. 5) demonstrates that the proposed method recovers a better high-resolution

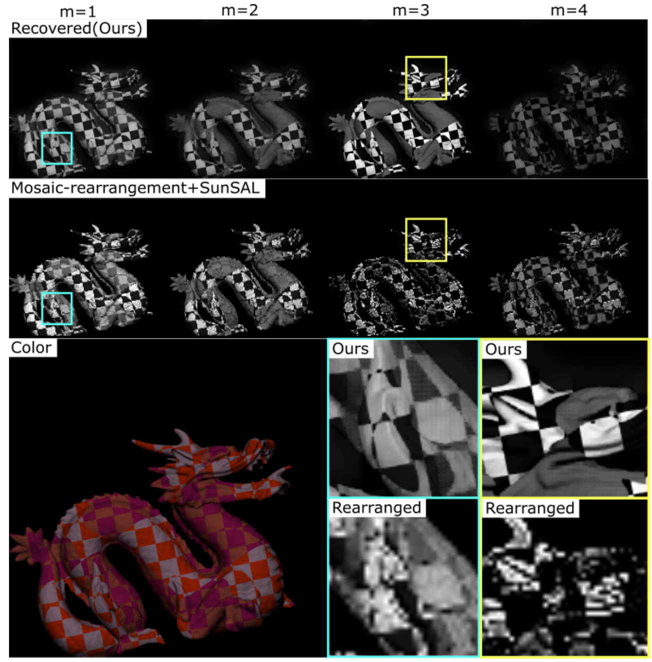

Fig. 2. Recovered high-resolution abundance maps (top, m is the index of the corresponding basis spectra or the endmember), low-resolution abundance maps (middle), color reference (bottom-left), and two zoomed-up regions of those abundance maps (bottom-right) for a simulated Dragon object textured with four Macbeth colors(-2, -7, -17, and -20).

multispectral image. The top-left is the band-1 pickups from the mosaic image, the rearranged low-resolution image, and the top-right is the completed high-resolution band-1 image. The bottom three rows show the rearranged low-resolution but zoomed-up textures, our recovered ones, and their color references. Color references are captured with iPhone-X, and there is some displacement to compare but not much. Sharp edges are retrieved in our completed results, and detailed spots of colorants are more clearly observed in our results than rearranged images.

Finally, in Fig. 6, we compare abundance maps that our method recovered, rearranged low-resolution maps, and TNTV [22] that performed the second-best among existing completion methods followed by the standard unmixing with VCA [23]. Here, we reasonably assume that estimated abundance maps under the extracted basis spectra seems roughly similar between rearranged and recovered maps using SUnSAL [33]. In this perspective, our recovered abundance maps looks similar to rearranged low-resolution maps, while the maps obtained after TNTV [22] differs. The contrast of each abundance map is clearly different between TNTV+SUnSAL and the mosaic-rearranged result. The edges of estimated maps are corrupted in rearranged low-resolution maps, while are clearly observed in our recovered maps. Distributions are different in TNTV [22], which we think is due to spectral bias in TNN formulation as described in the Introduction.

IV. Conclusion

We proposed a snapshot multispectral unmixing method with completing missing values of observed sparse tensors. Based on the basis spectra that are estimated with VCA, we recovered the full sensor size abundance maps under the nonnegative constraint and the anisotropic total variation regularization. We solved the proposed optimization problem by ADMM without any balancing parameter. A simulation experiment shows that our method performed better than the state-of-the-art tensor completion methods. We also demonstrate in a real data experiment that our method recovered the most reasonable, full sensor size abundance maps from a snapshot multispectral image.

References

[1] José M. Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gade, and Jocelyn Chanussot, Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol 5, 2, 2012.

[2] Stefan Livens, B. Delauré, Andy Lambrechts, and Nicolaas Tack, Hyperspectral imager development using direct deposition of interference filters, *Proceedings of the Small Satellites Systems and Services Symposium*, 2014

[3] Calum William, George S. D. Gordon, Timothy D. Wilkinson, and Sarah E. Bohndiek, Grayscale-to-color: Single-step fabrication of bespoke multispectral filter arrays, *ACS Photonics*, 6(12):3132–3141, 2019.

[4] Stefan Heist, Chen Zhang, Karl Reichwald, Peter Kühmstedt, Gunther Notni, and Andreas Tünnermann, 5D hyperspectral imaging: Fast and accurate measurement of surface shape and spectral characteristics using structured light, *Optics Express*, 26(18):23366–23379, 2018.

[5] Takeshi Uemori, Atsushi Ito, Yusuke Moriuchi, Alexander Gatto, and Jun Murayama, Skin-based identification from multispectral image data using CNNs, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*:12341–12350, 2019.

[6] Qinghua He and Ruikang K. Wang, Analysis of skin morphological features and real-time monitoring using snapshot hyperspectral imaging, *Journal of Biomedical Optics*, 10(11):5625–5638, 2019.

[7] Sofiane Mihoubi, Snapshot multispectral image demosaicing and classification, *Computer Science*, 2018.

[8] Lidan Miao, Hairong Qi, Rajeev Ramanath, and Wesley E. Snyder, Binary treebased generic demosaicking algorithm for multispectral filter arrays, *IEEE Transactions on Image Processing (TIP)*, 15(11):3550–3558, 2006.

[9] Yusuke Monno, Daisuke Kiku, Sunao Kikuchi, Masayuki Tanaka, and Masatoshi Okutomi, Multispectral Demosaicking with Novel Guide Image Generation and Residual Interpolation, *IEEE International Conference on Image Processing (ICIP)*:645–649, 2014.

[10] Kazuma Shinoda, Shu Ogawa, Yudai Yanagi, Madoka Hasegawa, Shigeo Kato, Masahiro Ishikawa, Hideki Komagata, and Naoki Kobayashi, Multispectral filter array and demosaicking for pathological images, *APSIPA ASC*, pp. 697–703, 2015.

[11] Kazuma Shinoda, Yasuo Ohtera, and Madoka Hasegawa, Snapshot multispectral polarization imaging using a photonic crystal filter array, *Optics Express* 26(12):15948–15961, 2018.

[12] Silvia Gandy, Benjamin Recht, and Isao Yamada, Tensor completion and low-n-rank tensor recovery via convex optimization, *Inverse Problems* 27(2), 2011.

[13] Teng-Yu Ji, Ting-Zhu Huang, Xi-Le Zhao, Tian-Hui Ma, and Gang Liu, Tensor completion using total variation and low-rank matrix factorization, *Information Sciences*, 326:243–257, 2016.

[14] Longhao Yuan, Chao Li, Danilo Mandic, Jianting Cao, and Qibin Zhao, Tensor ring decomposition with rank minimization on latent space: an efficient approach for tensor completion, *Association for the Advancement of Artificial Intelligence(AAAI)*, 2019.

[15] Misha E Kilmer, Karen Braman, Ning Hao, and Randy C. Hoover, Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging, *SIAM Journal on Matrix Analysis and Applications*, 34(1):148–172, 2013.

[16] Misha E Kilmer, Carla D Martin, and Lisa Perrone, A third- order generalization of the matrix svd as a product of third order tensors, *Tech. Report Technical Report TR–2008–4, Tufts University, Department of Computer Science*, 2008.

[17] Zemin Zhang, Gregory Ely, Shuchin Aeron, Ning Hao, and Misha Kilmer, Novel methods for multilinear data completion and de-noising based on tensor–svd, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*:3842–3849, 2014.

[18] Eric Kernfeld, Misha Kilmer, and Shuchin Aeron, Tensor–tensor products with invertible linear transforms, *Linear Algebra and its Applications*, 485:545–570, 2015.

[19] Canyi Lu, Jiashi Feng, Yudong Chen, Wei Liu, Zhouchen 958 Lin, and Shuicheng Yan, Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*:5249–5257, 2016.

[20] Canyi Lu, Jiashi Feng, Zhouchen Lin, and Shuicheng Yan, Exact low tubal rank tensor recovery from gaussian measurements, *International Joint Conferences on Artificial Intelligence Organization (IJCAI)*:2504–2510, 2018.

[21] Tai-Xiang Jiang, Michael K. Ng, Xi-Le Zhao, and Ting-Zhu Huang, Framelet representation of tensor nuclear norm for third-order tensor completion, 2020 *IEEE Transactions on Image Processing (TIP)*, 29:7233–7244, 2020.

[22] Duo Qiu, Minru Bai, Michael K. Ng, and Xiongjun Zhang, Robust low-rank tensor completion via transformed tensor nuclear norm with total variation regularization, *Neurocomputing*, 435:197–215, 2021.

[23] José M. P. Nascimento and Jose M. Bioucas-Dias, Vertex component analysis: A fast algorithm to unmix hyperspectral data, *IEEE Transactions on Geoscience and Remote Sensing*, 43(2):898–910, 2005.

[24] Emmanuel J. Candes, Xiaodong Li, Yi Ma, and John Wright, Robust principal component analysis?, *Journal of Association for Computing Machinery (ACM)*, 58(3), 2011.

[25] Wei He, Hongyan Zhang, and Liangpei Zhang, Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing, *IEEE Transactions on Geoscience and Remote Sensing*, 55(7):3909–3921, 2017.

[26] Marian-Daniel Iordache, José M. Bioucas–Dias, and Antonio Plaza, Total variation spatial regularization for sparse hyperspectral unmixing, *IEEE Transactions on Geoscience and Remote Sensing*, 50(11):4484–4502, 2012.

[27] Mia Rizkini and Masahiro Okuda, Local abundance regularization for hyperspectral sparse unmixing, *Asia–Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2016.

[28] Jing Qin, Harlin Lee, Jocelyn T. Chi, Lucas Drumetz, Jocelyn Chanussot, Yifei Lou, and Andrea L. Bertozzi, Blind hyperspectral unmixing based on graph total variation regularization, *IEEE Transactions on Geoscience and Remote Sensing*, 59(4):3338–3351, 2021.

[29] Amir Beck and Marc Teboulle, Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems, *IEEE Transactions on Image Processing (TIP)*, 18(11):2419–2434, 2009.

[30] Leonid I. Rudin, Stanley Osher, and Emad Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[31] Chunlin Wu and Xue-Cheng Tai, Augmented Lagrangian Method, Dual Methods, and Split Bregman Iteration for ROF, Vectorial TV, and High Order Models, *SIAM Journal on Imaging Sciences*, 3(3):300-339, 2010.

[32] Patrick Louis Combettes and Valérie R. Wajs, Signal recovery by proximal forward–backward splitting, *SIAM Journal on Multiscale Modeling and Simulation*, 4(4), 2005.

[33] J. Bioucas-Dias and M. Figueiredo, Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing, *2nd IEEE GRSS Workshop on Hyperspectral Image and SignalProcessing–WHISPERS*, 2010.

[34] Yilun Wang, Wotao Yin, and Yin Zhang, A fast algorithm for image deblurring with total variation regularization, *CAAM Technical Report TR07-10, Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77005,*, 2007.

[35] The Stanford 3D Scanning Repository. *at* http://graphics.stanford.edu/data/3Dscanrep/

[36] Matt Pharr and Greg Humphreys, Physically based rendering: From theory to implementation, *Morgan Kaufmann*, 2004.

TABLE I
COMPARISON RESULTS OF RECONSTRUCTION ERROR FOR SIMULATED MULTISPECTRAL IMAGES. B(D)-NO. DENOTES BUNNY OR DRAGON GEOMETRY AND THE REFLECTANCE OF SELECTED TEXTURED COLORS FROM MACBETH COLOR CHECKER.

	LRTC-TV		TRLRF		DCTNN		FTNN		TNTV		Ours	
Material	RelE	PSNR	RelE	PSNR	RelE	PSNR	RelE	PSNR	RelE	PSNR	RelE	PSNR
B-3-12	0.8469	18.43	1.237	14.20	0.4532	24.13	0.6791	20.73	0.3313	25.14	**0.0528**	**42.28**
B-5-9	0.8563	18.39	1.189	14.39	0.4493	23.98	0.6912	20.36	0.3230	25.04	**0.0460**	**43.30**
B-7-18	0.8403	20.66	1.383	14.01	0.4625	25.91	0.6764	22.44	0.4659	24.76	**0.0797**	**41.90**
D-4-16	0.8309	23.22	2.031	13.60	0.4531	28.55	0.7808	24.39	0.4212	27.20	**0.1911**	**36.37**
D-6-11	0.8436	21.50	1.802	13.37	0.4579	27.17	0.6524	23.86	0.4204	26.37	**0.0799**	**42.61**
D-8-15	0.8413	20.84	1.750	13.20	0.4588	26.45	0.6681	22.98	0.3906	26.28	**0.1014**	**39.11**
B-3-9-12-14	0.8288	20.64	1.507	14.16	0.4561	25.74	0.6467	22.55	0.3296	26.72	**0.0790**	**39.12**
B-2-7-13-17	0.8406	18.87	1.192	14.42	0.4446	24.34	0.7867	19.93	0.3272	24.54	**0.0880**	**36.41**
B-1-5-11-21	0.8375	20.14	1.481	13.92	0.4399	25.74	0.7202	21.53	0.2803	27.31	**0.0867**	**38.68**
D-2-7-17-20	0.8417	21.28	1.789	13.66	0.4590	26.66	0.6022	23.57	0.3758	26.38	**0.1044**	**39.15**
D-5-8-12-14	0.8371	20.94	1.740	13.56	0.4611	26.45	0.6483	23.18	0.3922	26.05	**0.1046**	**38.64**
D-4-10-15-19	0.8449	23.08	2.071	14.65	0.4665	28.30	0.5021	26.12	0.3992	28.04	**0.1120**	**39.99**

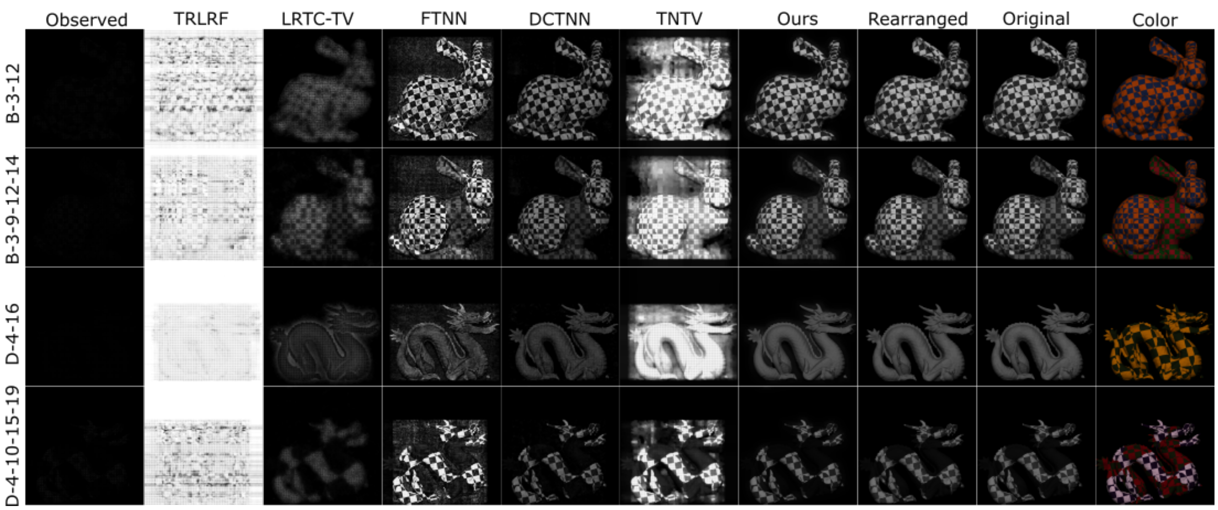

Fig. 3. Simulation comparison between completion results for band-1 images between our method and state-of-the-art completion methods. Observed band-1 image, rearranged low-resolution band-1 image, and color image are also shown for reference.

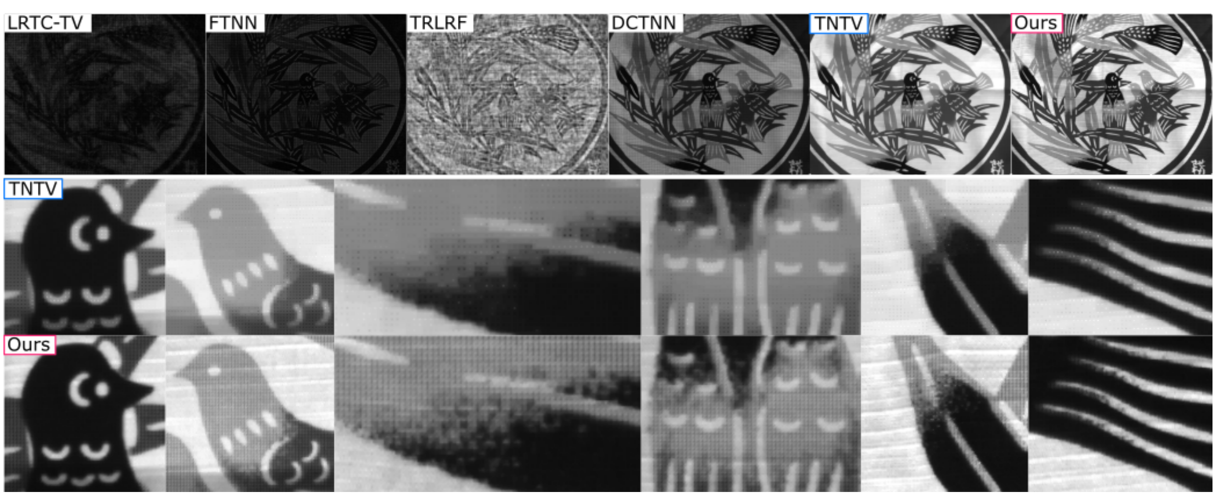

Fig. 4. Real data comparison between completion results for band-1 images of our method and state-of-the-art tensor completion methods (first row); enlarged views from TNTV of second-best quality (second row) and ours of the best (third row).

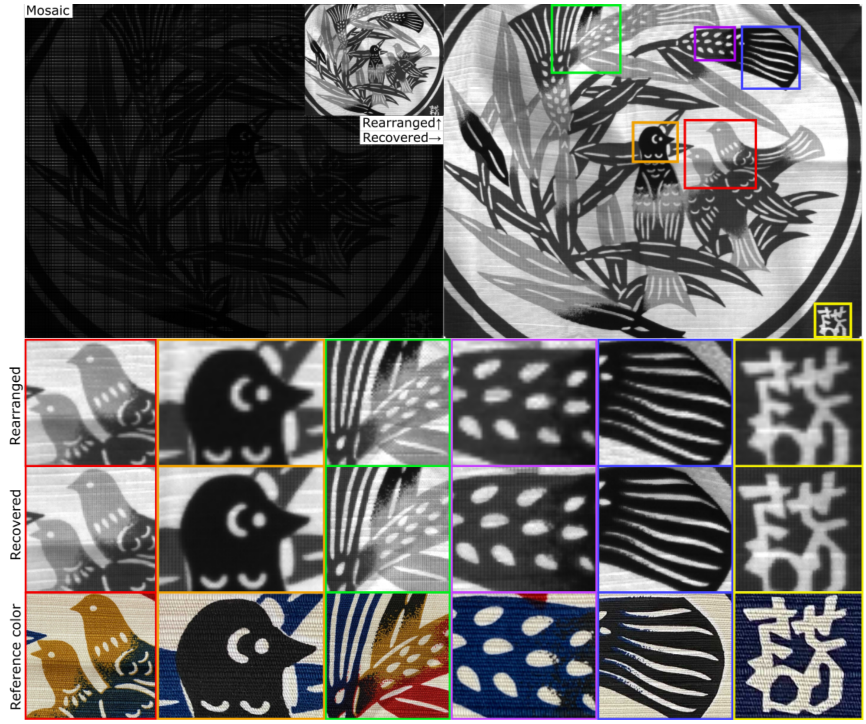

Fig. 5. Band-1 of the mosaic, rearranged, and our completed image. Bottom three rows show zoomed-up textures of selected regions with reference colors.

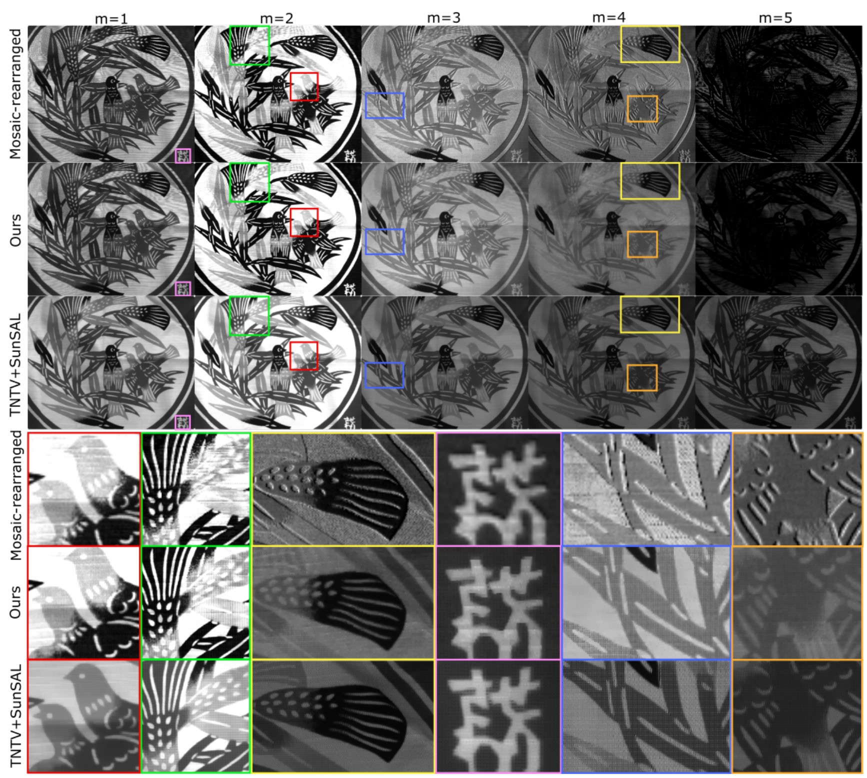

Fig. 6. The estimated abundance maps (m is the index of the corresponding basis spectra or the endmember), rearranged low–resolution result, and TNTV [22] with unmixing result. Bottom three rows are zoomed-up selected regions.