FAST FUSION OF HYPERSPECTRAL AND MULTISPECTRAL IMAGES: A TUCKER APPROXIMATION APPROACH

Clémence Prévost, Pierre Chainais, Remy Boyer

FAST FUSION OF HYPERSPECTRAL AND MULTISPECTRAL IMAGES : A TUCKER APPROXIMATION APPROACH

C. Prévost, P. Chainais and R. Boyer

Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France

ABSTRACT

Hyperspectral super-resolution based on coupled Tucker decomposition has been recently considered in the remote sensing community. The state-of-the-art approaches did not fully exploit the coupling information contained in hyperspectral and multispectral images of the same scene. In this paper, we propose a new algorithm that overcomes this limitation. It accounts for both the high-resolution and the low-resolution information in the model, by solving a set of least-squares problems. In addition, we provide exact recovery conditions for the super-resolution image in the noiseless case. Our simulations show that the proposed algorithm achieves good reconstruction with low complexity.

***Index Terms*—** hyperspectral super-resolution, data fusion, low-rank tensor factorizations, recovery, least-squares problem

1. INTRODUCTION

Hyperspectral devices produce hyperspectral images (HSI) with high spectral resolution. The compromise between signal-to-noise ratio, spatial and spectral resolutions force the HSIs to have low spatial resolution [1]. On the other hand, multispectral images (MSI) have high spatial resolution, but a restricted number of spectral bands. The hyperspectral super-resolution (HSR) problem [2] recovers a super-resolution image (SRI) with both high spatial and high spectral resolutions from co-registered HSI and MSI of the same scene.

Matrix-based approaches to HSR [3, 4, 5, 6] performed a coupled low-rank factorization of the matricized HSI and MSI. More recently, tensor approaches were envisioned, motivated by the 3-dimensional structure of the observations, and possible uniqueness guarantees offered by tensor low-rank factorizations. The works of [7, 8] were the first to consider tensor-based HSR and paved the way for various works, see, e.g., [9, 10, 11, 12].

The Tucker decomposition was steadily considered [13, 9, 12] for HSR. In [9], two algorithms based on the Singular Value Decomposition (SVD) were proposed [9]. The first one, named Super-resolution based on COupled Tucker Tensor approximation (SCOTT), recovered the Tucker factors based on the SVD followed by solving a least-squares problem. The second one assumed that degradation between the HSI and SRI was unknown, therefore it was referred to as BSCOTT (Blind-SCOTT). The Tucker model in [9] was extended to account for localized changes in [12].

These algorithms had a major limitation: they did not fully exploit the coupling information between the observations. The blind algorithm BSCOTT only accounted for a portion of the degradation. In contrast, SCOTT was non-blind but was highly sub-optimal. Indeed, the SVD step only extracted high-resolution information from one single observation. Hence the low-resolution information available from the observation model was totally ignored. Since no iterations were performed, the solution to the least squares problem did not incorporate the coupling information to the Tucker factors.

In this paper, we propose a new Tucker-based algorithm that overcomes the limitations of SCOTT and BSCOTT. It fully accounts for the coupling constraints in the HSR model. We propose a simple way to incorporate additional information to the Tucker factors. The SVD is performed on both observations and is followed by solving a least-squared problem. We prove that exact recovery of the SRI is still achievable with the proposed approach. Our experiments on real datasets show that the proposed approach achieves good reconstruction performance with low complexity.

Notation. We follow the notations of [14, 15]. We use the symbols $\boxtimes$ and $\odot$ for the Kronecker and the Khatri-Rao products. We use $\mathrm{vec}\{\cdot\}$ for the standard column-major vectorization of a tensor or a matrix. Operation $\bullet_p$ denotes contraction on the pth index of a tensor; for instance, for a tensor $\boldsymbol{\mathcal{A}}$ and a matrix $\mathbf{M}$, $[\boldsymbol{\mathcal{A}} \bullet_1 \mathbf{M}]_{\ell jk} = \sum_\ell \mathcal{A}_{ijk} M_{i\ell}$. For a tensor $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{I\times J\times K}$, its unfoldings are $\mathbf{Y}^{(1)} \in \mathbb{R}^{JK\times I}$, $\mathbf{Y}^{(2)} \in \mathbb{R}^{IK\times J}$ and $\mathbf{Y}^{(3)} \in \mathbb{R}^{IJ\times K}$. We use shorthand notation for the multilinear product: $[\![\boldsymbol{\mathcal{G}}; \mathbf{U}, \mathbf{V}, \mathbf{W}]\!] = \boldsymbol{\mathcal{G}} \bullet_1 \mathbf{U} \bullet_2 \mathbf{V} \bullet_3 \mathbf{W}$. It can be obtained from the Higher-Order SVD (HOSVD). If $R_1 = \mathrm{rank}(\mathbf{Y}^{(1)})$, $R_2 = \mathrm{rank}(\mathbf{Y}^{(2)})$ and $R_3 = \mathrm{rank}(\mathbf{Y}^{(3)})$, then the multilinear product is called Tucker decomposition of $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{I\times J\times K}$ and $(R_1, R_2, R_3) < (I, J, K)$ are the multilinear ranks.

2. HYPERSPECTRAL SUPER-RESOLUTION

2.1. Problem statement and observation model

We consider an MSI $\boldsymbol{\mathcal{Y}}_M \in \mathbb{R}^{I\times J\times K_M}$ and HSI $\boldsymbol{\mathcal{Y}}_H \in \mathbb{R}^{I_H\times J_H\times K}$ acquired from sensors (for instance, LANDSAT or QuickBird). The MSI has higher spatial resolution than the HSI ($I_H < I, J_H < J$), but lower spectral resolution ($K_M < K$). The $\boldsymbol{\mathcal{Y}}_M$ and $\boldsymbol{\mathcal{Y}}_H$ are viewed as two degraded versions of a single SRI $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{I\times J\times K}$. The HSR problem consists in recovering $\boldsymbol{\mathcal{Y}}$ from $\boldsymbol{\mathcal{Y}}_M$ and $\boldsymbol{\mathcal{Y}}_H$. As commonly adopted in the literature [7, 9, 10], we consider the following degradation model:

$$\begin{cases} \boldsymbol{\mathcal{Y}}_H &= \boldsymbol{\mathcal{Y}} \bullet_1 \mathbf{P}_1 \bullet_2 \mathbf{P}_2 + \boldsymbol{\mathcal{E}}_H, \\ \boldsymbol{\mathcal{Y}}_M &= \boldsymbol{\mathcal{Y}} \bullet_3 \mathbf{P}_3 + \boldsymbol{\mathcal{E}}_M, \end{cases} \tag{1}$$

where $\boldsymbol{\mathcal{E}}_M$ and $\boldsymbol{\mathcal{E}}_H$ are noise terms with independent and identically distributed entries, having zero mean and variances σ_H^2 and σ_M^2, respectively. The matrix $\mathbf{P}_3 \in \mathbb{R}^{K_M\times K}$ is the spectral degradation matrix, and $\mathbf{P}_1 \in \mathbb{R}^{I_H\times I}$, $\mathbf{P}_2 \in \mathbb{R}^{J_H\times J}$ are the spatial degradation matrices, *i.e.*, we assume (for simplicity) that the spatial degradation is separable. This assumption is reasonable thanks to the commonly accepted Wald's protocol [16], that uses isotropic Gaussian blurring and downsampling.

This work was partly supported by the ANR project "Chaire IA Sherlock" ANR-20-CHIA-0031-01 hold by P. Chainais, as well as by the national support within the *programme d'investissements d'avenir* ANR-16-IDEX-0004 ULNE and Région HDF.

Similarly to [9], we utilize a Tucker-based coupled model. Assume that the SRI $\boldsymbol{\mathcal{Y}}$ admits a Tucker decomposition with given multilinear ranks (R_1, R_2, R_3) as

$$\boldsymbol{\mathcal{Y}} = [\![\boldsymbol{\mathcal{G}}; \mathbf{U}, \mathbf{V}, \mathbf{W}]\!], \tag{2}$$

where $\mathbf{U} \in \mathbb{R}^{I \times R_1}$, $\mathbf{V} \in \mathbb{R}^{J \times R_2}$ and $\mathbf{W} \in \mathbb{R}^{K \times R_3}$ are the factor matrices and $\boldsymbol{\mathcal{G}} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$ is the core tensor. Model (1) becomes

$$\begin{cases} \boldsymbol{\mathcal{Y}}_H &= [\![\boldsymbol{\mathcal{G}}; \mathbf{P}_1\mathbf{U}, \mathbf{P}_2\mathbf{V}, \mathbf{W}]\!] + \boldsymbol{\mathcal{E}}_H, \\ \boldsymbol{\mathcal{Y}}_M &= [\![\boldsymbol{\mathcal{G}}; \mathbf{U}, \mathbf{V}, \mathbf{P}_3\mathbf{W}]\!] + \boldsymbol{\mathcal{E}}_M. \end{cases} \tag{3}$$

The aim of HSR is to recover the factors $\mathbf{U},\ \mathbf{V},\ \mathbf{W}$ and $\boldsymbol{\mathcal{G}}$ from (3).

2.2. Tucker-based HSR: state-of-the-art and its limitations

In [9], a Tucker-based algorithm called SCOTT was proposed. It consisted in three simple steps. First, $\widehat{\mathbf{U}},\ \widehat{\mathbf{V}},\ \widehat{\mathbf{W}}$ were recovered as the dominant right singular vectors of the unfoldings $\mathbf{Y}_M^{(1)},\ \mathbf{Y}_M^{(2)}$ and $\mathbf{Y}_H^{(3)}$, respectively. Then, the tensor $\widehat{\boldsymbol{\mathcal{G}}}$ was reconstructed as:

$$\begin{aligned} \underset{\boldsymbol{\mathcal{G}}}{\operatorname{argmin}}\ f_T(\boldsymbol{\mathcal{G}}, \widehat{\mathbf{U}}, \widehat{\mathbf{V}}, \widehat{\mathbf{W}}) = \lambda \|\boldsymbol{\mathcal{Y}}_H - [\![\boldsymbol{\mathcal{G}}; \mathbf{P}_1\widehat{\mathbf{U}}, \mathbf{P}_2\widehat{\mathbf{V}}, \widehat{\mathbf{W}}]\!]\|_F^2 \\ + \|\boldsymbol{\mathcal{Y}}_M - [\![\boldsymbol{\mathcal{G}}; \widehat{\mathbf{U}}, \widehat{\mathbf{V}}, \mathbf{P}_3\widehat{\mathbf{W}}]\!]\|_F^2, \end{aligned} \tag{4}$$

where $\lambda = \frac{\sigma_M^2}{\sigma_H^2}$. This step consisted in minimizing the quadradic cost function:

$$\left\| \underbrace{\begin{bmatrix} \sqrt{\lambda}\widehat{\mathbf{W}} \boxtimes \mathbf{P}_2\widehat{\mathbf{V}} \boxtimes \mathbf{P}_1\widehat{\mathbf{U}} \\ \mathbf{P}_M\widehat{\mathbf{W}} \boxtimes \widehat{\mathbf{V}} \boxtimes \widehat{\mathbf{U}} \end{bmatrix}}_{\mathbf{X}} \operatorname{vec}\{\widehat{\boldsymbol{\mathcal{G}}}\} - \underbrace{\begin{bmatrix} \sqrt{\lambda}\operatorname{vec}\{\boldsymbol{\mathcal{Y}}_H\} \\ \operatorname{vec}\{\boldsymbol{\mathcal{Y}}_M\} \end{bmatrix}}_{\boldsymbol{z}} \right\|_F^2,$$

which could be solved through normal equations and yielded

$$\operatorname{vec}\{\widehat{\boldsymbol{\mathcal{G}}}\} = \left(\mathbf{X}^\mathsf{T}\mathbf{X}\right)^{-1}\mathbf{X}^\mathsf{T}\boldsymbol{z}, \tag{5}$$

see [9] for more details. Finally, the low-rank approximation of the SRI was reconstructed using (2). In [9], experiments showed the good performance of SCOTT with a low computational complexity.

However, SCOTT suffered from major limitations. First, the estimation of the factor matrices was sub-optimal. For instance, $\widehat{\mathbf{U}}$ was recovered only from $\boldsymbol{\mathcal{Y}}_M$, hence the low-resolution spatial information contained in $\boldsymbol{\mathcal{Y}}_H$ was totally ignored. Recovery of $\widehat{\mathbf{V}}$ and $\widehat{\mathbf{W}}$ suffered from the same drawbacks. In other words, the second step of SCOTT did not incorporate at all the low-resolution information. Hence SCOTT did not consider all the information from (3).

2.3. Proposed approach

In this paper, we propose an improved Tucker-based algorithm that overcomes the limitations of SCOTT. The new estimation strategy accounts for both the high-resolution and the low-resolution information available in model (1). This strategy was considered in [8] for estimation of $\mathbf{W}$ only.

In model (1), high-resolution information is always directly available in at least one observation. While $\boldsymbol{\mathcal{Y}}_H$ contains high-resolution spectral information, $\boldsymbol{\mathcal{Y}}_M$ contains high-resolution spatial information. Therefore, one observation can always be used as reference for recovery of $\widehat{\mathbf{U}},\ \widehat{\mathbf{V}},\ \widehat{\mathbf{W}}$. The proposed approach will be applied for separate recovery of these matrices. First, we will utilize the reference observation to obtain a subspace describing the high-resolution information. We will also extract information from the other tensor even if its content has been degraded by the observation model. Finally, we will reconstruct the low-rank factor as a tradeoff between the two sources of information. For each matrix, we will consider tensor unfoldings corresponding to the dimension of interest. In the following, we will explain how to obtain $\widehat{\mathbf{W}}$. The same strategy holds accordingly for $\widehat{\mathbf{U}}$ and $\widehat{\mathbf{V}}$.

We wish to recover a low-rank estimate $\widehat{\boldsymbol{\mathcal{Y}}}$ of the SRI such that

$$\widehat{\mathbf{Y}}^{(3)} = \mathbf{D}\widehat{\mathbf{W}}^\mathsf{T}, \tag{6}$$

where $\mathbf{D} \in \mathbb{R}^{IJ \times R_3}$ and $\widehat{\mathbf{W}} \in \mathbb{R}^{K \times R_3}$ is the factor matrix containing high-resolution spectral information.

In practice, high-resolution spectral information is directly available from HSI $\boldsymbol{\mathcal{Y}}_H$. It can be obtained from the SVD with rank $R_3 < K$, which gives the solution to the low-rank approximation problem of $\mathbf{Y}_H^{(3)}$ in the least-squared sense [17]. This operation produces a tensor $\boldsymbol{\mathcal{Z}}_H \in \mathbb{R}^{I_H \times J_H \times K}$ such that

$$\mathbf{Z}_H^{(3)} = \mathbf{L}_H\boldsymbol{\Sigma}_H\mathbf{W}_H^\mathsf{T}, \tag{7}$$

where $\mathbf{W}_H \in \mathbb{R}^{K \times R_3}$ contains the dominant right singular vectors of $\mathbf{Z}_H^{(3)}$. Its columns describe the subspace containing high-resolution information [18]. We want to ensure the coherence between $\mathbf{W}_H$ and $\widehat{\mathbf{W}}$. In other words, $\widehat{\mathbf{W}}$ needs to satisfy

$$\widehat{\mathbf{W}}^\mathsf{T} = \arg\min_{\widehat{\mathbf{W}}^\mathsf{T}} \|\widehat{\mathbf{W}}^\mathsf{T} - \mathbf{W}_H^\mathsf{T}\|_F^2. \tag{8}$$

In our approach, we also want to exploit the degraded spectral information contained in $\boldsymbol{\mathcal{Y}}_M$. Under the assumption that $R_3 \leq K_M$, this information can be obtained similarly to (7) by denoising of $\mathbf{Y}_M^{(3)}$ and yields a tensor $\boldsymbol{\mathcal{Z}}_M \in \mathbb{R}^{I \times J \times K_M}$ such that

$$\mathbf{Z}_M^{(3)} = \mathbf{L}_M\boldsymbol{\Sigma}_M\mathbf{W}_M^\mathsf{T}, \tag{9}$$

where $\mathbf{W}_M \in \mathbb{R}^{K_M \times R_3}$ describes the subspace containing spectral information degraded by $\mathbf{P}_3$.

Coherence between degraded information and high-resolution information projected onto the span of $\mathbf{W}_H$ leads $\widehat{\mathbf{W}}$ to satisfy

$$\widehat{\mathbf{W}}^\mathsf{T} = \arg\min_{\widehat{\mathbf{W}}^\mathsf{T}} \|\widehat{\mathbf{W}}^\mathsf{T}\mathbf{W}_H\mathbf{W}_H^\mathsf{T}\mathbf{P}_3^\mathsf{T} - \mathbf{W}_M^\mathsf{T}\|_F^2. \tag{10}$$

Finally, we wish to recover $\widehat{\mathbf{W}}$ as a tradeoff between the two sources of information. By combining (8) and (10), we want that

$$\widehat{\mathbf{W}}^\mathsf{T} = \arg\min_{\widehat{\mathbf{W}}^\mathsf{T}} \|\widehat{\mathbf{W}}^\mathsf{T} \underbrace{\mathbf{W}_H\mathbf{W}_H^\mathsf{T}\mathbf{P}_3^\mathsf{T}}_{\mathbf{A}} - \mathbf{W}_M^\mathsf{T}\|_F^2 + \lambda\|\widehat{\mathbf{W}}^\mathsf{T} - \mathbf{W}_H^\mathsf{T}\|_F^2, \tag{11}$$

which can be obtained as

$$\widehat{\mathbf{W}}^\mathsf{T} = \left(\mathbf{W}_M^\mathsf{T}\mathbf{A}^\mathsf{T} + \lambda\mathbf{W}_H^\mathsf{T}\right)\left(\mathbf{A}\mathbf{A}^\mathsf{T} + \lambda\mathbf{I}\right)^{-1}. \tag{12}$$

The matrices $\widehat{\mathbf{U}}$ and $\widehat{\mathbf{V}}$ can be recovered similarly as

$$\widehat{\mathbf{U}}^\mathsf{T} = \left(\lambda\mathbf{U}_H^\mathsf{T}\mathbf{B}^\mathsf{T} + \mathbf{U}_M^\mathsf{T}\right)\left(\lambda\mathbf{B}\mathbf{B}^\mathsf{T} + \mathbf{I}\right)^{-1}, \tag{13}$$

$$\widehat{\mathbf{V}}^\mathsf{T} = \left(\lambda\mathbf{V}_H^\mathsf{T}\mathbf{C}^\mathsf{T} + \mathbf{V}_M^\mathsf{T}\right)\left(\lambda\mathbf{C}\mathbf{C}^\mathsf{T} + \mathbf{I}\right)^{-1}, \tag{14}$$

Algorithm 1: Proposed approach
input : $\mathcal{Y}_M, \mathcal{Y}_H, R_1, R_2, R_3$
output: $\widehat{\mathcal{Y}} \in \mathbb{R}^{I\times J\times K}$
1. $\mathbf{U}_H, \mathbf{V}_H, \mathbf{W}_H \leftarrow$HOSVD$(\mathcal{Y}_H)$, $\mathbf{U}_M, \mathbf{V}_M, \mathbf{W}_M \leftarrow$HOSVD$(\mathcal{Y}_M)$;
2. $\widehat{\mathbf{U}}^\mathsf{T} \leftarrow$ (13), $\widehat{\mathbf{V}}^\mathsf{T} \leftarrow$ (14), $\widehat{\mathbf{W}}^\mathsf{T} \leftarrow$ (12);
3. $\widehat{\mathcal{G}} \leftarrow \underset{\mathcal{G}}{\operatorname{argmin}}\ f_T(\mathcal{G}, \widehat{\mathbf{U}}, \widehat{\mathbf{V}}, \widehat{\mathbf{W}})$;
4. $\widehat{\mathcal{Y}} = [\![\widehat{\mathcal{G}}; \widehat{\mathbf{U}}, \widehat{\mathbf{V}}, \widehat{\mathbf{W}}]\!]$.

by setting $\mathbf{B} = \mathbf{U}_M\mathbf{U}_M^\mathsf{T}\mathbf{P}_1^\mathsf{T}$ and $\mathbf{C} = \mathbf{V}_M\mathbf{V}_M^\mathsf{T}\mathbf{P}_2^\mathsf{T}$, respectively.

The proposed approach is summarized in Algorithm 1.

The ranks (R_1, R_2, R_3) cannot exceed (I_H, J_H, K_M).As suggested in [8], we circumvent this limitation by applying Algorithm 1 to corresponding non-overlapping subblocks of the MSI and the HSI. This is based on the hypothesis that smaller blocks in the observations are more likely to contain a small number of materials and spatial features [19]. Thus, $\mathcal{Y}_H$ and $\mathcal{Y}_M$ are divided into corresponding $L \times L$ subblocks of spatial dimensions $\frac{I_H}{L} \times \frac{J_H}{L}$ and $\frac{I}{L} \times \frac{J}{L}$. This strategy is summarized in Algorithm 2.

Algorithm 2: Block version of Algorithm 1
input : $\mathcal{Y}_M, \mathcal{Y}_H, R_1, R_2, R_3$
Split $\mathcal{Y}_H$ and $\mathcal{Y}_M$ into corresponding subblocks;
for $\ell = 1, \ldots, L$ **do**
  Apply Algorithm 1 to the pair of blocks of $\mathcal{Y}_H$ and $\mathcal{Y}_M$, and reconstruct the corresponding block of $\widehat{\mathcal{Y}}$.
end

The total computational complexity of Algorithm 1 is

- $\mathcal{O}(\min(R_1, R_2, R_3)(IJK_M + I_HJ_HK))$ flops for the truncated SVDs;
- $\mathcal{O}(\min(R_3^3 + (R_1R_2)^3; R_1^3 + (R_2R_3)^3))$ flops for solving the Sylvester equation in Step 4;

and it is dominated by the cost of the truncated SVD. It is smaller than that of SCOTT, which required $\mathcal{O}(\min(R_1, R_2)IJK_M + R_3I_HJ_HK)$ flops for recovery of the factor matrices.

3. RECOVERABILITY OF THE COUPLED MODEL

In this subsection, we give uniqueness results for the SRI tensor recovery by the approach of Algorithm 2.

3.1. Deterministic recovery

Theorem 3.1. *Let a Tucker decomposition of $\mathcal{Y}$ with multilinear ranks (R_1, R_2, R_3) be $\mathcal{Y} = [\![\mathcal{G}; \mathbf{U}, \mathbf{V}, \mathbf{W}]\!]$, where $\mathcal{G} \in \mathbb{R}^{R_1\times R_2\times R_3}$, $\mathbf{U} \in \mathbb{R}^{I\times R_1}$, $\mathbf{V} \in \mathbb{R}^{J\times R_2}$, $\mathbf{W} \in \mathbb{R}^{K\times R_3}$ have full column rank. We also assume that there is no noise,* i.e.*, $\mathcal{E}_H, \mathcal{E}_M = \mathbf{0}$ in* (1)*. If*

$$\operatorname{rank}\left(\mathbf{Y}_M^{(1)}\right) = R_1,\ \operatorname{rank}\left(\mathbf{Y}_M^{(2)}\right) = R_2,\ \operatorname{rank}\left(\mathbf{Y}_H^{(3)}\right) = R_3, \quad (15)$$

and if

$$\operatorname{rank}(\mathbf{P}_1\mathbf{U}) = R_1,\ \operatorname{rank}(\mathbf{P}_2\mathbf{V}) = R_2,\ \operatorname{rank}(\mathbf{P}_3\mathbf{W}) = R_3, \quad (16)$$

then Algorithm 1 recovers $\mathcal{Y}$ correctly, i.e.*, there exists only one $\widehat{\mathcal{Y}}$ with multilinear ranks (R_1, R_2, R_3) such that $\widehat{\mathcal{Y}} \bullet_1 \mathbf{P}_1 \bullet_2 \mathbf{P}_2 = \mathcal{Y}_H$ and $\widehat{\mathcal{Y}} \bullet_3 \mathbf{P}_3 = \mathcal{Y}_M$.*

Conversely, if

$$\operatorname{rank}(\mathbf{P}_1\mathbf{U})\operatorname{rank}(\mathbf{P}_2\mathbf{U}) < R_1R_2 \text{ and } \operatorname{rank}(\mathbf{P}_3\mathbf{W}) < R_3 \quad (17)$$

then there exists infinitely many $\widehat{\mathcal{Y}}$ of the form $\widehat{\mathcal{Y}} = [\![\widehat{\mathcal{G}}; \widehat{\mathbf{U}}, \widehat{\mathbf{V}}, \widehat{\mathbf{W}}]\!]$ such that $\widehat{\mathcal{Y}} \bullet_1 \mathbf{P}_1 \bullet_2 \mathbf{P}_2 = \mathcal{Y}_H$ and $\widehat{\mathcal{Y}} \bullet_3 \mathbf{P}_3 = \mathcal{Y}_M$.

3.2. Generic recovery

Now, we formulate recovery guarantees in the generic case, *i.e.*, that hold for a random tensors with multilinear ranks (R_1, R_2, R_3).

Theorem 3.2. *Assume that $\mathbf{P}_1 \in \mathbb{R}^{I_H\times I}$, $\mathbf{P}_2 \in \mathbb{R}^{J_H\times J}$, and $\mathbf{P}_3 \in \mathbb{R}^{K_M\times K}$ are fixed full row-rank matrices. Let*

$$\mathcal{Y} = [\![\mathcal{G}; \mathbf{U}, \mathbf{V}, \mathbf{W}]\!],$$

where $\mathcal{G} \in \mathbb{R}^{R_1\times R_2\times R_3}$, $R_1 \le I$, $R_2 \le J$, $R_3 \le K$, and $\mathbf{U} \in \mathbb{R}^{I\times R_1}$, $\mathbf{V} \in \mathbb{R}^{J\times R_2}$, $\mathbf{W} \in \mathbb{R}^{K\times R_3}$ are random matrices, distributed according to an absolutely continuous probability distribution. We also assume that $\mathcal{E}_M, \mathcal{E}_H = \mathbf{0}$ in (1)*.*

If $(R_1, R_2, R_3) \le (I_H, J_H, K_M)$, and

$$R_1 \le R_2R_3,\ R_2 \le R_1R_3 \text{ and } R_3 \le R_1R_2, \quad (18)$$

then with probability 1 *there exists a unique tensor $\widehat{\mathcal{Y}}$ such that $\widehat{\mathcal{Y}}_M = \mathcal{Y}_M$ and $\widehat{\mathcal{Y}}_H = \mathcal{Y}_H$ that can be recovered by Algorithm 1.*

Proof. First, without loss of generality, we can replace $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$ with the following of same size (see [9, Section V.B]):

$$\widetilde{\mathbf{P}}_1 = \begin{bmatrix}\mathbf{I}_{I_H}\\ \mathbf{0}\end{bmatrix}^\mathsf{T},\ \widetilde{\mathbf{P}}_2 = \begin{bmatrix}\mathbf{I}_{J_H}\\ \mathbf{0}\end{bmatrix}^\mathsf{T},\ \widetilde{\mathbf{P}}_3 = \begin{bmatrix}\mathbf{I}_{K_M}\\ \mathbf{0}\end{bmatrix}^\mathsf{T}. \quad (19)$$

Therefore, under the assumptions on distribution of $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ the following implications hold with probability 1,

$$\begin{aligned} R_1 \le I_H &\Rightarrow \operatorname{rank}(\mathbf{U}_{1:I_H,:}) = R_1,\\ R_2 \le J_H &\Rightarrow \operatorname{rank}(\mathbf{V}_{1:J_H,:}) = R_2,\\ R_3 \le K_M &\Rightarrow \operatorname{rank}(\mathbf{W}_{1:K_M,:}) = R_3. \end{aligned}$$

Next, we are going to show how (18) imply (15). We will prove it only for the first condition (the others are analogous). Note that the first unfolding can be written as

$$\mathbf{Y}_M^{(1)} = (\mathbf{W}_{1:K_M,:} \boxtimes \mathbf{V})\mathbf{G}^{(1)}\mathbf{U}^\mathsf{T}. \quad (20)$$

Due to the dimensions of the terms in the product, this matrix in (20) is at most rank R_1. Due to semicontinuity of the rank function, $\mathbf{Y}_M^{(1)}$ will be generically of rank R_1 if we can provide just a single example of $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$, $\mathcal{G}$, achieving the condition $\operatorname{rank}(\mathbf{Y}_M^{(1)}) = \operatorname{rank}(\mathbf{U}) = R_1$. Due to the dimension in (20), such an example can be found under the condition $R_1 \le \min(R_3, K_M)R_2$. It is given by

$$\mathbf{U} = \begin{bmatrix}\mathbf{I}_{R_1}\\ \mathbf{0}\end{bmatrix}, \mathbf{V} = \begin{bmatrix}\mathbf{I}_{R_2}\\ \mathbf{0}\end{bmatrix}, \mathbf{W} = \begin{bmatrix}\mathbf{I}_{R_3}\\ \mathbf{0}\end{bmatrix}, \mathbf{G}^{(1)} = \begin{bmatrix}\mathbf{I}_{R_1}\\ \mathbf{0}\end{bmatrix},$$

which completes the proof. □

Remark 3.3. *Our conditions for exact recovery are formulated in the noiseless case. This assumption can be justified by the fact that real hyperspectral and multispectral images tend to have a high signal-to-noise ratio. In practice, Theorem 3.2 guarantees exact reconstruction under some* small *additive noise, hence the ranks must be chosen accordingly, as it will be discussed in Section 4.1.*

4. EXPERIMENTS

The performance was evaluated using metrics used in [20], including Reconstruction Signal-to-Noise Ratio (R-SNR), Cross-Correlation (CC), Spectral Angle Mapper (SAM) and Relative dimensionless Global Error in Synthesis (ERGAS). The runtime was evaluated using the `tic` and `toc` functions of MATLAB.

The matrices $\mathbf{P}_1, \mathbf{P}_2$ were generated following Wald's protocol [16] with a downsampling ratio $d = 2$ and a Gaussian kernel of size $q = 9$. The matrix $\mathbf{P}_3$ contained the spectral response functions of the Sentinel-2 instrument[1]. The HSI $\mathcal{Y}_H$ was obtained by degradation of the true SRI by $\mathbf{P}_1$ and $\mathbf{P}_2$ and the MSI $\mathcal{Y}_M$ was obtained by spectral degradation by $\mathbf{P}_3$. The spectral bands of $\mathcal{Y}_H$ and $\mathcal{Y}_M$ were normalized and the true SRI was denoised [21]. We compared our algorithm to matrix-based approaches: HySure [5], CNMF [3] and FUSE [23]. We also considered tensor methods: STEREO [7] for CP decomposition, SCOTT [9] for Tucker and CB-STAR for block-term decomposition [12]. We chose the ranks according to [12].

4.1. Choice of the ranks

We considered a portion of the Lockwood dataset with $\mathcal{Y} \in \mathbb{R}^{88\times 88\times 173}$. We studied the impact of the multilinear ranks on the performance of Algorithm 2. We considered a $[2,2]$-block pattern and a $[4,4]$-block pattern. For each scenario, we computed the R-SNR as a function of $R_1 = R_2$ and R_3 for ranks satisfying Theorem 3.2. The results were displayed in Figure 1.

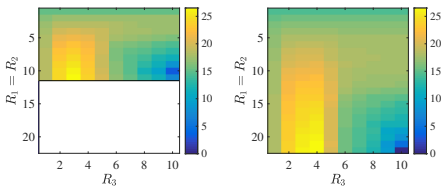

Fig. 1. R-SNR as a function of $R_1 = R_2$ and R_3, $[4,4]$ (left) and $[2,2]$-block pattern (right), Lockwood.

High R-SNR was achieved even if the rank conditions were very restrictive. The best R-SNR was obtained for $R_3 = 3$ (resp. $R_3 = 4$), and large $R_1 = R_2$. A drop of performance could be observed for larger R_3. Indeed, selecting high R_3 would extract least significant singular values, leading the reconstruction to be corrupted by noise. In the following, the ranks were chosen according to Figure 1 so that they maximized the R-SNR.

4.2. Reconstruction performance

The first dataset was Isabella Lake with $\mathcal{Y} \in \mathbb{R}^{88\times 120\times 173}$. White Gaussian noise was added to the HSI and MSI to yield 30dB SNR each, thus $\lambda = 1$. For Algorithm 2, we considered ranks $(11, 11, 5)$ and $(22, 22, 5)$ for the $[4,4]$ and $[2,2]$ splitting scenarios, respectively. The results were available in Table 1 and the two best results of each columns were shown in bold[2]. In Figure 2, we showed false color plots of the reconstructed images. The RGB color space of the reference SRI was used.

The CP-based algorithm STEREO and Algorithm 2 yielded the best metrics. Our proposed approach performed slightly better than SCOTT. Algorithm 2 with $[4,4]$ blocks had the lowest run time.

We then considered the Lockwood dataset. White Gaussian noise was added to the HSI and MSI to yield 40dB and 10dB SNR, respectively, thus we had $\lambda = 1 \cdot 10^{-6}$. We took $[4,4]$ and $[2,2]$-block patterns and we chose the ranks $(11, 11, 3)$ and $(22, 22, 4)$, respectively. The results were shown in Table 2 and Figure 3.

[1] Available for download online.

[2] The numbers between brackets for Algorithm 2 denoted the number of blocks.

Algorithm	R-SNR	CC	SAM	ERGAS	Time (s)
SCOTT	25.360	0.9842	3.5026	9.3365	2.6874
Alg. 3 $[4,4]$	27.318	**0.9891**	**2.4733**	**6.6369**	**0.4384**
Alg. 3 $[2,2]$	**27.373**	**0.9895**	2.7335	**6.9332**	1.5724
STEREO [7]	**27.870**	0.9885	3.0376	9.5818	2.2347
CB-STAR [12]	26.787	0.9881	2.9085	7.7640	13.671
CNMF [3]	24.975	0.9882	**2.5072**	8.3609	6.6395
HySure [5]	22.073	0.9697	5.3823	18.208	13.353
FUSE [23]	24.505	0.9846	3.0626	7.6932	**0.4650**

Table 1. Reconstruction metrics, Isabella Lake.

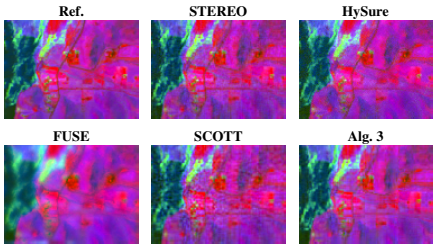

Fig. 2. False color plots of reconstructed SRI, Isabella Lake.

Algorithm	R-SNR	CC	SAM	ERGAS	Time (s)
SCOTT	1.1814	0.4432	32.640	170.89	0.5112
Alg. 3 $[4,4]$	**17.722**	**0.8641**	4.8926	13.614	**0.1834**
Alg. 3 $[2,2]$	16.321	0.8016	6.6554	19.108	0.7048
STEREO [7]	10.531	0.6456	15.635	50.665	2.0948
CB-STAR [12]	13.664	0.7528	10.052	26.174	44.966
CNMF [3]	16.617	0.8252	4.6671	27.234	2.5084
HySure [5]	17.565	0.8472	**4.5703**	**12.735**	7.3825
FUSE [23]	**17.610**	**0.8716**	**3.2267**	**10.388**	**0.3455**

Table 2. Reconstruction metrics, Lockwood.

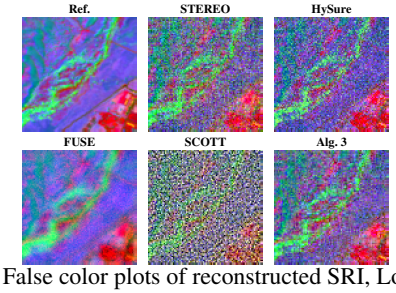

Fig. 3. False color plots of reconstructed SRI, Lockwood.

The best metrics were provided by FUSE and Algorithm 2. SCOTT had the worst performance since it relied only on the (very noisy) MSI observation for estimation of $\widehat{\mathbf{U}}$ and $\widehat{\mathbf{V}}$. Algorithm 2 had runtime comparable to that of SCOTT.

5. CONCLUSION

We proposed a new Tucked-based algorithm for HSR. It is able to fully exploit both high-resolution and low-resolution information contained in the model. We showed that it reaches good performance at a low complexity. In particular, it performs better than state-of-the-art in scenarios when one observation is very noisy, with a competitive runtime. Its good performance naturally raises the question of its statistical efficiency: this matter will be addressed in future works.

6. REFERENCES

[1] G. A. Shaw and H. K. Burke, "Spectral imaging for remote sensing," *Lincoln laboratory journal*, vol. 14, no. 1, pp. 3–28, 2003.

[2] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Trans. Geosci. Remote Sens.*, vol. 5, no. 2, pp. 29–56, 2017.

[3] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, 2012.

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

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

[6] Q. Wei, J. M. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Multiband image fusion based on spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7236–7249, 2016.

[7] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Hyperspectral Super-Resolution: A Coupled Tensor Factorization Approach," *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6503–6517, 2018.

[8] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Hyperspectral Super-Resolution: Combining Low Rank Tensor and Matrix Structure," in *2018 IEEE ICIP*, Oct. 2018, pp. 3318–3322.

[9] C. Prévost, K. Usevich, P. Comon, and D. Brie, "Hyperspectral Super-Resolution with Coupled Tucker Approximation: Identifiability and SVD-based algorithms," *IEEE Trans. Signal Process.*, vol. 68, pp. 931–946, 2020.

[10] G. Zhang, X. Fu, K. Huang, and J. Wang, "Hyperspectral super-resolution: A coupled nonnegative block-term tensor decomposition approach," in *2019 IEEE CAMSAP*, 2019, Guadeloupe, West Indies.

[11] M. Ding, X. Fu, T.-Z. Huang, J. Wang, and X.-L. Zhao, "Hyperspectral super-resolution via interpretable block-term tensor modeling," *arXiv e-prints*, p. arXiv:2006.10248, June 2020.

[12] R. A. Borsoi, C. Prévost, K. Usevich, D. Brie, J. M. Bermudez, and C. Richard, "Coupled tensor decomposition for hyperspectral and multispectral image fusion with inter-image variability," *IEEE J. Sel. Topics Signal Process.*, 2021.

[13] R. Dian, L. Fang, and S. Li, "Hyperspectral image super-resolution via non-local sparse tensor factorization," *IEEE Conf. on Comput. Vision and Pattern Recogn. (CVPR)*, pp. 5344– 5353, 2017.

[14] P. Comon, "Tensor Decompositions, State of the Art and Applications," in *Mathematics in Signal Processing V*, J. G. McWhirter and I. K. Proudler, Eds., pp. 1–24. Clarendon Press, Oxford, UK, 2002.

[15] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

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

[17] G. H. Golub, A. Hoffman, and G. W. Stewart, "A generalization of the eckart-young-mirsky matrix approximation theorem," *Linear Algebra and its applications*, vol. 88, pp. 317–327, 1987.

[18] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3551–3582, 2017.

[19] T. Imbiriba, R. A. Borsoi, and J. M. Bermudez, "Generalized linear mixing model accounting for endmember variability," in *2018 IEEE ICASSP*, 2018, pp. 1862–1866.

[20] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, M. Selva, and C. Chen, "25 years of pansharpening: a critical review and new developments," *Signal and Image Process. for Remote Sens.*, pp. 533–548, 2011.

[21] R. E. Roger and J. F. Arnold, "Reliably estimating the noise in AVIRIS hyperspectral images," *International Journal of Remote Sensing*, vol. 17, no. 10, pp. 1951–1962, 1996.

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

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