

COUPLED TENSOR LOW-RANK MULTILINEAR APPROXIMATION FOR HYPERSPECTRAL SUPER-RESOLUTION

C. Prévost¹, K. Usevich¹, P. Comon², D. Brie¹

¹ CRAN, Université de Lorraine, CNRS, Vandoeuvre-lès-Nancy, France

² GIPSA-Lab, Univ. Grenoble Alpes, CNRS, Saint-Martin d'Hères, France
clemence.prevast@univ-lorraine.fr, konstantin.usevich@univ-lorraine.fr
pierre.comon@gipsa-lab.grenoble-inp.fr, david.brie@univ-lorraine.fr

ABSTRACT

We propose a novel approach for hyperspectral super-resolution that is based on low-rank tensor approximation for a coupled low-rank multilinear (Tucker) model. We show that the correct recovery holds for a wide range of multilinear ranks. For coupled tensor approximation, we propose an SVD-based algorithm that is simple and fast, but with a performance comparable to that of the state-of-the-art methods.

Index Terms— hyperspectral super-resolution, data fusion, low-rank tensor factorizations, recovery, identifiability

1. INTRODUCTION

The problem of *hyperspectral super-resolution* (HSR) [1] has recently attracted much interest from the signal processing community. It consists in fusing a multispectral image (MSI), which has a good spatial resolution but few spectral bands, and a hyperspectral image (HSI), whose spatial resolution is lower than that of MSI. The aim of this method is to recover a *super-resolution image* (SRI), which possesses both good spatial and spectral resolutions. This problem is closely related to hyperspectral pansharpening [2, 3], where HSI is fused with a panchromatic image.

Many methods were developed for the HSR problem: CNMF [4], methods based on solving a Sylvester equation [5], HySure [6], FUMI [7], to name a few. Motivated by the linear mixing model widely used in hyperspectral image unmixing, most of these methods are based on a coupled low-rank factorization of the matricized HSI and MSI.

Recently, a promising tensor-based method that makes use of the inherent 3D nature of HSI was proposed [8], where the HSR is reformulated as a coupled CP (canonical polyadic) approximation. An alternating least squares algorithm is proposed, achieving reconstruction performance that is competitive with the state of the art. The key property underlying the approach of [8] is that the coupled CP decomposition is identifiable. This approach was also successfully used recently for a super-resolution problem in medical imaging [9]. Still, it has several drawbacks: for instance, the appropriate rank of the CP decomposition is not known *a priori* and may be

unrelated to the number of endmembers; the rank can also be large, which may affect the computational complexity and convergence of the ALS iterations. A blind block-based version was proposed in [10] to overcome some of these issues.

In this paper, we propose an approach which is similar in spirit to [8], but which is based on another type of low-rank tensor factorization: we reformulate the HSR problem as a coupled multilinear (Tucker) approximation. First, we propose a simple closed-form algorithm that is similar to the multilinear (or higher-order) SVD. Second, we show that, although the Tucker decomposition is not identifiable¹, the correct recovery of the SRI holds for a wide range of ranks, in contrast to the ordinary Tucker factorization that is generally not unique. Our experiments show that the proposed algorithm has a performance comparable to the one of [8]. Finally, we show that the proposed approach is applicable in the case of hyperspectral pansharpening (unlike [8], which requires the MSI to have at least two spectral bands).

This paper is organized as follows. In Section 2, we recall the HSR problem and the STEREO algorithm proposed in [8]. Section 3 contains our proposed coupled Tucker model and an SVD-based algorithm for tensor approximation (SCOTT). In Section 4 we give our main identifiability result (recoverability) for the coupled Tucker model. Section 5 contains the numerical experiments. Proofs and details omitted in this paper can be found in the extended version of the paper [11].

Notation. In this paper we mainly follow [12] in what concerns the tensor notation (see also [13]). We use the symbol \boxtimes for the Kronecker product, and \odot for the Khatri-Rao product. We use $\text{vec}\{\cdot\}$ for the standard column-major vectorization of a tensor or a matrix. Operation \bullet_p denotes contraction on the p th index of a tensor; for instance, $[\mathbf{A} \bullet_1 \mathbf{M}]_{ijk} = \sum_\ell \mathbf{A}_{\ell jk} M_{i\ell}$. For a tensor \mathcal{G} and matrices \mathbf{U} , \mathbf{V} and \mathbf{W} , the following shorthand notation is used

$$[\mathcal{G}; \mathbf{U}, \mathbf{V}, \mathbf{W}] = \mathcal{G} \bullet_1 \mathbf{U} \bullet_2 \mathbf{V} \bullet_3 \mathbf{W}.$$

For matrices $\mathbf{A} \in \mathbb{R}^{I \times F}$, $\mathbf{B} \in \mathbb{R}^{J \times F}$, $\mathbf{C} \in \mathbb{R}^{K \times F}$, we will use a shorthand notation for the polyadic decomposition

$$[\mathbf{A}, \mathbf{B}, \mathbf{C}] = [\mathcal{I}_F; \mathbf{A}, \mathbf{B}, \mathbf{C}]$$

This work was partially supported by the Agence Nationale de la Recherche under grant OPTIFIN (ANR-15-CE10-0007).

¹This is the reason why the Tucker model was discarded in [8] as a potential model for hyperspectral super-resolution.

where $\mathcal{I}_F \in \mathbb{R}^{F \times F \times F}$ is a diagonal tensor of ones. For a tensor $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$, its first unfolding is denoted by $\mathbf{Y}^{(1)} \in \mathbb{R}^{J \times K \times I}$. By $\text{tSVD}_R(\mathbf{X})$ we denote a matrix containing R leading right singular vectors of the matrix \mathbf{X} .

2. HYPERSPECTRAL DATA FUSION PROBLEM

2.1. Problem statement and degradation model

We consider an MSI cube $\mathcal{Y}_M \in \mathbb{R}^{I \times J \times K_M}$ and a HSI cube $\mathcal{Y}_H \in \mathbb{R}^{I_H \times J_H \times K}$ acquired from existing sensors (for instance, LANDSAT or QuickBird). The acquired MSI and HSI represent the same target, and \mathcal{Y}_M and \mathcal{Y}_H are viewed as two degraded versions of a single SRI data cube $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$. The hyperspectral data fusion problem [1] consists in recovering \mathcal{Y} from \mathcal{Y}_M and \mathcal{Y}_H .

In this paper, we adopt the following degradation model, written as contraction of SRI with degradation matrices:

$$\begin{cases} \mathcal{Y}_M = \mathcal{Y} \bullet_3 \mathbf{P}_M + \mathcal{E}_M, \\ \mathcal{Y}_H = \mathcal{Y} \bullet_1 \mathbf{P}_1 \bullet_2 \mathbf{P}_2 + \mathcal{E}_H, \end{cases} \quad (1)$$

where \mathcal{E}_M and \mathcal{E}_H denote the noise terms, $\mathbf{P}_M \in \mathbb{R}^{K_M \times K}$ is the spectral degradation matrices (for example, a selection-averaging matrix), $K_M < K$, and $\mathbf{P}_1 \in \mathbb{R}^{I_H \times I}$, $\mathbf{P}_2 \in \mathbb{R}^{J_H \times J}$, $I_H < I$, $J_H < J$, are the spatial degradation matrices, i.e. we assume (for simplicity) that the spatial degradation is separable; this is a valid assumption, for example, for the commonly accepted Wald's protocol [14], which uses Gaussian blurring and downsampling in both spatial dimensions. In this paper we consider only the case when the degradation matrices $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_M$ are known.

2.2. CP-based approach (STEREO)

In [8] it was proposed to model the SRI as a tensor with low CP rank, i.e. $\mathcal{Y} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]$, where $\mathbf{A} \in \mathbb{R}^{I \times F}$, $\mathbf{B} \in \mathbb{R}^{J \times F}$, $\mathbf{C} \in \mathbb{R}^{K \times F}$ are the factor matrices of the CPD and F is the CP rank. In this case, the HSR problem can be formulated as

$$\underset{\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}}{\text{minimize}} \quad f_{CP}(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}), \quad (2)$$

where $f_{CP}(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}) =$

$$\|\mathcal{Y}_H - [\mathbf{P}_1 \hat{\mathbf{A}}, \mathbf{P}_2 \hat{\mathbf{B}}, \hat{\mathbf{C}}]\|_F^2 + \lambda \|\mathcal{Y}_M - [\hat{\mathbf{A}}, \hat{\mathbf{B}}, \mathbf{P}_M \hat{\mathbf{C}}]\|_F^2,$$

which is a coupled CP approximation problem. For the case when there is no noise $\mathcal{E}_H, \mathcal{E}_M = \mathbf{0}$, the coupled CP model is (generically) identifiable if $F \leq \min\{2^{\lfloor \log_2(K_M J) \rfloor - 2}, I_H J_H\}$, see [8] for more details.

To solve (2), an alternating least squares (ALS) optimization algorithm was proposed in [8], called STEREO (Super-Resolution Tensor Reconstruction), which cyclicly updates $\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}$ while minimizing $f_{CP}(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}})$. The initial guess is chosen using the rank- F CP approximation of \mathcal{Y}_M , and solving one least squares problem. The updates in the ALS iterations can be computed by using efficient solvers for the (generalized) Sylvester equation [15], [16]. See [8] or [11] for more details.

3. TUCKER-BASED DATA FUSION

3.1. Model and approximation problem

In this paper, we propose a Tucker-based coupled model as an alternative to STEREO. Let $\mathbf{R} = (R_1, R_2, R_3)$ be the multi-linear ranks of the SRI \mathcal{Y} , and let $\mathcal{Y} = [\mathcal{G}; \mathbf{U}, \mathbf{V}, \mathbf{W}]$ be its Tucker decomposition of the SRI, where $\mathbf{U} \in \mathbb{R}^{I \times R_1}$, $\mathbf{V} \in \mathbb{R}^{J \times R_2}$ and $\mathbf{C} \in \mathbb{R}^{K \times R_3}$ are the factor matrices of the tensor and $\mathcal{G} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$ is the core tensor.

With these notations, Equation (1) becomes

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

The HSR formulation is thus

$$\underset{\hat{\mathcal{G}}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{\mathbf{W}}}{\text{minimize}} \quad f_T(\hat{\mathcal{G}}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{\mathbf{W}}), \quad \text{where} \quad (3)$$

$$\begin{aligned} f_T(\hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{\mathbf{W}}, \hat{\mathcal{G}}) = & \|\mathcal{Y}_H - [\hat{\mathcal{G}}; \mathbf{P}_1 \hat{\mathbf{U}}, \mathbf{P}_2 \hat{\mathbf{V}}, \hat{\mathbf{W}}]\|_F^2 \\ & + \lambda \|\mathcal{Y}_M - [\hat{\mathcal{G}}; \hat{\mathbf{U}}, \hat{\mathbf{V}}, \mathbf{P}_M \hat{\mathbf{W}}]\|_F^2. \end{aligned} \quad (4)$$

3.2. An SVD-based algorithm

A suboptimal solution to problem (3) can be found by an HOSVD-like Algorithm 1, named SCOTT (Superresolution based on COupled Tucker Tensor approximation).

input : $\mathcal{Y}_M \in \mathbb{R}^{I \times J \times K_M}, \mathcal{Y}_H \in \mathbb{R}^{I_H \times J_H \times K}$,
 R_1, R_2, R_3
output: $\hat{\mathcal{Y}} \in \mathbb{R}^{I \times J \times K}$
 $\hat{\mathbf{U}} \leftarrow \text{tSVD}_{R_1}(\mathbf{Y}_M^{(1)}), \hat{\mathbf{V}} \leftarrow \text{tSVD}_{R_2}(\mathbf{Y}_M^{(2)}),$
 $\hat{\mathbf{W}} \leftarrow \text{tSVD}_{R_3}(\mathbf{Y}_H^{(3)}),$
 $\hat{\mathcal{G}} \leftarrow \underset{\mathcal{G}}{\text{argmin}} \quad f_T(\mathcal{G}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{\mathbf{W}})$
 $\hat{\mathcal{Y}} = [\hat{\mathcal{G}}; \hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{\mathbf{W}}].$

Algorithm 1: SCOTT

The computation of $\hat{\mathcal{G}}$ in Algorithm 1 is a least squares problem which can be solved through normal equations, which in turn can be viewed as a Sylvester equation; hence efficient solvers [15], [16] can be used. As shown in [11], the computational complexity of Algorithm 1 is comparable to that of STEREO.

4. RECOVERABILITY OF THE COUPLED MODEL

In this subsection, we give a generic uniqueness result for the tensor recovery in the coupled Tucker model.

Theorem 4.1. Assume that $\mathbf{P}_1 \in \mathbb{R}^{I_H \times I}$, $\mathbf{P}_2 \in \mathbb{R}^{J_H \times J}$, and $\mathbf{P}_M \in \mathbb{R}^{K_M \times K}$ are fixed full row-rank matrices. Let

$$\mathcal{Y} = [\mathcal{G}; \mathbf{U}, \mathbf{V}, \mathbf{W}], \quad (5)$$

where $\mathcal{G} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$, $R_1 \leq I$, $R_2 \leq J$, $R_3 \leq K$, 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).
1. If $R_3 \leq K_M$ or $(R_1, R_2) \leq (I_H, J_H)$, and

$$\begin{cases} R_1 \leq \min(R_3, K_M)R_2, \\ R_2 \leq \min(R_3, K_M)R_1 \\ R_3 \leq \min(R_1, I_H) \min(R_2, J_H), \end{cases} \quad (6)$$

then with probability 1 there exists a unique tensor $\hat{\mathcal{Y}}$ such that $\hat{\mathcal{Y}}_M = \mathcal{Y}_M$ and $\hat{\mathcal{Y}}_H = \mathcal{Y}_H$.

2. If $R_3 > K_M$ and $(R_1 > I_H$ or $R_2 > J_H)$, then the reconstruction is non-unique, i.e. there exist an continuum of $\hat{\mathcal{Y}}$ such that $\hat{\mathcal{Y}}_M = \mathcal{Y}_M$ and $\hat{\mathcal{Y}}_H = \mathcal{Y}_H$; in fact, $\|\hat{\mathcal{Y}} - \mathcal{Y}\|$ can be arbitrary large.

The proof of Theorem 4.1 is given in [11]. We illustrate the statement of Theorem 4.1 for the case $I = J$, $I_H = J_H$ and $R_1 = R_2$; in Figure 1 we show that the space of parameters (R_1, R_3) is split into two regions: recoverable and non-recoverable. The hatched area corresponds to the parameters where condition (6) is not satisfied.

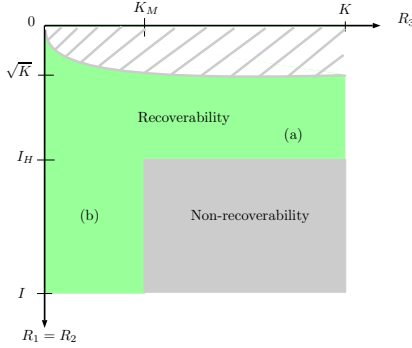


Fig. 1. recoverability region depending on R_1 and R_3

In [11], two other interesting observations are made. First, it is shown that SRI recoverability is related to correct recovery in tensor completion. Second, it is shown that the case of unknown spatial degradation operators can be treated by a blind version of SCOTT.

5. EXPERIMENTS

All simulations were run on a MacBook Pro with 2.4 GHz Intel Core i5 and 8GB RAM. For basic tensor operations we used TensorLab 3.0 [17]. The results are reproducible and the MATLAB codes are available online at

https://github.com/cprevost4/HSR_Tucker.

More details on the experiments are available in the extended version of the paper [11].

As for the experimental setup, we follow [8] as closely as possible. The main performance metric that we use is *reconstruction Signal-to-Noise ratio* (R-SNR) introduced in [3]

$$\text{R-SNR} = 10 \log_{10} \left(\frac{\|\mathcal{Y}\|_F^2}{\|\hat{\mathcal{Y}} - \mathcal{Y}\|_F^2} \right).$$

In all the examples, as in [8], the bands corresponding to water absorption are removed. For all the experiments, the degradation matrices $\mathbf{P}_1, \mathbf{P}_2$ are generated following Wald's protocol, and the downsampling factor is chosen to be 4, see [14]. The matrix \mathbf{P}_M is a selection-averaging matrix that splits the spectral range into equal parts.

5.1. Choice of multilinear ranks

In this subsection, we have a closer look at the Indian Pines dataset, available online at [18], that was acquired by the AVIRIS sensor, with LANDSAT specifications used for spectral degradation. In this case, $\mathcal{Y} \in \mathbb{R}^{144 \times 144 \times 200}$, $\mathcal{Y}_M \in \mathbb{R}^{144 \times 144 \times 6}$ and $\mathcal{Y}_H \in \mathbb{R}^{36 \times 36 \times 200}$. We analyze the performance of SCOTT (R-SNR and the value of the cost function f_T defined in (4)) w.r.t. the multilinear rank employed in the Tucker decomposition.

Figure 2 shows the R-SNR and cost function value for $R_1 = R_2$ in $[10 : 50]$ and R_3 in $[2 : 25]$ for which the recoverability condition holds (see Section 4 and Fig. 1).

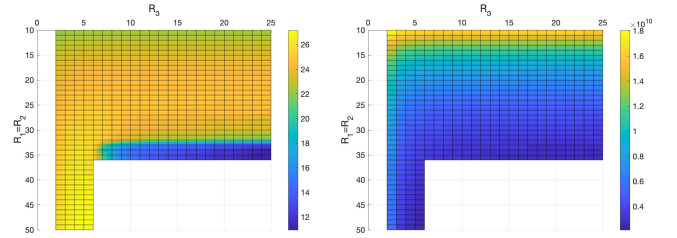


Fig. 2. R-SNR (left) and f_T (right) as functions of R_1 and R_3

While the cost function decreases as R_1 and R_3 increase, the best reconstruction error (given by R-SNR) is achieved in one of the two recoverability subregions in Fig. 1: (a) ($R_3 \geq K_M$ and $R_1 \leq I_H$) and (b) ($R_3 \leq K_M$ and $R_1 \geq I_H$). For subregion (b), it seems better to take $R_3 = K_M$ and R_1 as large as possible, while for subregion (a), we notice a drop of the R-SNR around $R_1 = I_H$.

We also performed the optimization of (3) with the structured data fusion framework implemented in Tensorlab [17]. In our experiment, for all the values in Fig. 2, the optimization procedure stopped after very few iterations and did not improve the cost function. This result seems reasonable, because it is known [19] that truncated HOSVD gives a very good solution for a low-rank Tucker approximation.

5.2. Comparison with other algorithms

In this subsection, we compare the performance of STEREO, SCOTT, and HySure [6]. In addition to R-SNR, different standard metrics from [3] are used (ERGAS, SAM, CC). We also show the computational time for each algorithm, given by the `tic` and `toc` functions of MATLAB. As in [8], we run STEREO for 10 iterations.

5.2.1. Semi-real data

First, we compare the results for some semi-real data, available online at [18]. The representative multilinear ranks

are chosen to be $[40, 40, 6]$, $[30, 30, 16]$ and $[24, 24, 25]$, which correspond to different regions of recoverability. For STEREO, we choose ranks $F = 50$ and $F = 100$ as in [8], and for HySure we use the number of groundtruth materials as the number of endmembers. Table 1 shows these metrics for the Indian Pines dataset. In general, SCOTT achieves results comparable to STEREO in the case $F = 50$ used in [8]. However, the case of STEREO $F = 100$ gives slightly better results than tensor rank $F = 50$, especially in terms of R-SNR. SCOTT appears to be fastest in this case (although we do not claim it, because the speed depends on the implementation). Fast algorithms for Sylvester equations are used in both cases.

In [11], we also evaluated performance of the algorithms for reconstruction of endmembers; our experiments show that SCOTT allows for accurate reconstruction of spectra.

Algorithm	R-SNR	CC	SAM	ERGAS	time
STEREO 50	26.93	0.9	2.24	1.02	3.06
STEREO 100	28.53	0.92	2.01	0.87	5.24
SCOTT (40,40,6)	26.32	0.89	2.34	1.07	3.03
SCOTT (30,30,16)	23.82	0.84	2.77	1.38	0.71
SCOTT (24,24,25)	24.65	0.87	2.56	1.23	0.31
HySure	26.96	0.9	2.35	1.01	16.09

Table 1. Comparison of algorithms, Indian Pines dataset

The second dataset we consider is the Salinas A-scene, also available at [18]: it consists in a portion of the bigger Salinas dataset. In this case, $\mathbf{Y} \in \mathbb{R}^{80 \times 84 \times 204}$, and $\mathbf{Y}_M \in \mathbb{R}^{80 \times 84 \times 6}$, $\mathbf{Y}_H \in \mathbb{R}^{20 \times 21 \times 204}$. The results are shown in Table 2. In this case, $R_3 = 6$ seems to be the best choice for SCOTT, which agrees with the number of materials. Our algorithm is again rather fast, except for the case where it shows a better performance than STEREO (with $F = 100$, which is the rank chosen in [8] for this dataset).

In Figures 3 and 4, we show a single slice of the reconstructed SRI for various algorithms and both datasets.

Algorithm	R-SNR	CC	SAM	ERGAS	time
STEREO 50	33.03	0.97	0.89	4.11	1.63
STEREO 100	32.83	0.93	0.63	6.99	2.71
SCOTT (40,40,6)	31.5	0.95	0.7	5.17	2.51
SCOTT (14,14,15)	24.97	0.92	1.76	5.5	0.06
SCOTT (10,15,25)	23.31	0.91	1.97	5.62	0.05
SCOTT (30,30,6)	30.06	0.94	0.93	5.26	0.58
SCOTT (58, 58, 6)	32.63	0.95	0.52	5.14	23.58
HySure	28.53	0.94	0.9	5.13	4.92

Table 2. Comparison of algorithms, Salinas A-scene dataset

5.2.2. Pansharpening problem

Here, we address the pansharpening problem, which consists in fusion of a hyperspectral image and a panchromatic image (PAN) \mathbf{Y}_P . PAN is obtained by averaging over the full spectral range of the groundtruth SRI, so that $\mathbf{P}_M \in \mathbb{R}^{1 \times K}$ and $\mathbf{Y}_P \in \mathbb{R}^{I \times J \times 1}$.

In this case, the STEREO algorithm is not applicable, since its initialization² is based on the CPD of the MSI (which

²We also tried different initializations for STEREO, including an initialization based on the SCOTT solution.

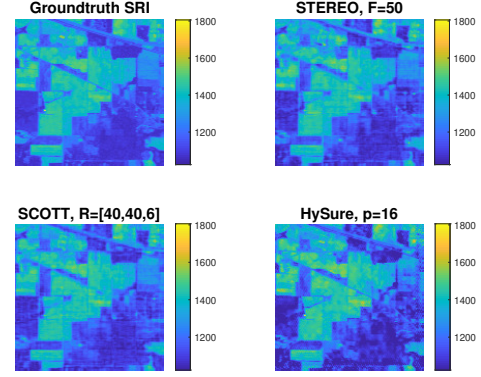


Fig. 3. Band 160, Indian Pines dataset

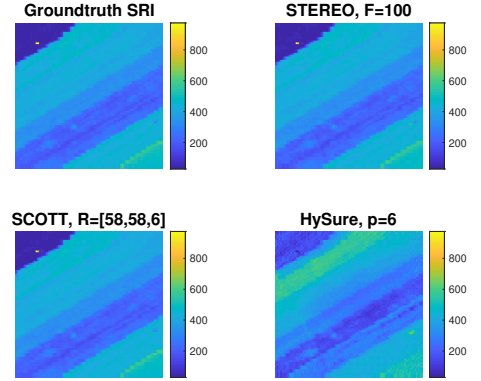


Fig. 4. Band 160, Salinas-A dataset

is a matrix in the case of PAN images). However, SCOTT and HySure can perform the data fusion.

In Table 3, the metrics are shown for different multilinear ranks and for the HySure method, for the Indian Pines dataset. In this case as well, the previous conclusions on the metrics hold, showing that the SCOTT algorithm is able to give a reasonable solution to the pansharpening problem as well.

Algorithm	R-SNR	CC	SAM	ERGAS	time
SCOTT (24, 24, 25)	20.59	0.78	4.36	1.93	0.3
SCOTT (30, 30, 16)	18.48	0.7	5.37	2.48	0.72
SCOTT (35, 35, 6)	11.38	0.41	10.53	5.71	2.14
HySure	18.01	0.62	5.97	2.68	18.97

Table 3. Metrics for different algorithms, Indian Pines dataset

6. CONCLUSION

In this paper, we proposed a novel coupled Tucker model for the HSR problem. We showed that the model is *recoverable*, and that a very simple SVD-based algorithm can be used for the super-resolution problem (including the case of pansharpening), giving results that are comparable with the CP-based approach. We hope that this work opens new perspectives on using various tensor factorizations for hyperspectral super-resolution. Still several interesting questions remain, for example, how to enlarge the recoverability range for the multilinear ranks.

7. REFERENCES

- [1] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 2, pp. 29–56, 2017.
- [2] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, J. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, and N. Yokoya, "Hyperspectral pansharpening: A review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 3, no. 3, pp. 27–46, 2015.
- [3] 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 Processing for Remote Sensing*, pp. 533–548, 2011.
- [4] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 528–537, 2012.
- [5] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3658–3668, 2015.
- [6] M. Simoões, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3373–3388, 2015.
- [7] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Multiband image fusion based on spectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7236–7249, 2016.
- [8] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Hyperspectral super-resolution: A coupled tensor factorization approach," *IEEE Transactions on Signal Processing*, vol. 66, no. 24, pp. 6503–6517, 2018.
- [9] J. Hatvani, A. Basarab, J.-Y. Tourneret, M. Gyöngy, and D. Kouamé, "A Tensor Factorization Method for 3D Super-Resolution with Application to Dental CT," *IEEE Transactions on Medical Imaging*, 2018, doi:10.1109/TMI.2018.2883517.
- [10] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W. Ma, "Hyperspectral super-resolution: Combining low rank tensor and matrix structure," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, Oct 2018, pp. 3318–3322.
- [11] C. Prévost, K. Usevich, P. Comon, and D. Brie, "Hyperspectral Super-Resolution with Coupled Tucker Approximation: Identifiability and SVD-based algorithms," Tech. Rep., 2018, available from <https://hal.archives-ouvertes.fr/hal-01911969>.
- [12] P. Comon, "Tensors: A brief introduction," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 44–53, 2014.
- [13] Tamara G. Kolda and Brett W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [14] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogrammetric engineering and remote sensing*, vol. 63, no. 6, pp. 691–699, 1997.
- [15] R.H. Bartels and G.W. Stewart, "Solution of the matrix equation $AX+XB=C$," *Communications of the ACM*, vol. 15, no. 9, pp. 820–826, 1972.
- [16] G. Golub, S. Nash, and C. Van Loan, "A Hessenberg-Schur method for the problem $AX+XB=C$," *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 909–913, 1979.
- [17] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," Mar. 2016, Available online.
- [18] "Grupo de inteligencia computacional (gic)," http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes, accessed: 2018-10-23.
- [19] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.