Tensor-based image fusion accounting for inter-image variability: Recoverability and algorithms

Ricardo BORSOI[1], Clémence PRÉVOST[2], Konstantin USEVICH[1], David BRIE[1], José BERMUDEZ[3], Cédric RICHARD[4]

[1]Univ. Lorraine, CNRS, UMR 7039 CRAN – nom.prenom[at]univ-lorraine[dot]fr

[2]Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL – nom.prenom[at]univ-lille[dot]fr

[3]Fed. Univ. Santa Catarina, Florianópolis, Brésil

[4]Univ. Côte d'Azur, CNRS, OCA, Lagrange – nom.prenom[at]unice[dot]fr

Résumé – Dans cet article, nous proposons une approche pour la reconstruction d'images à haute résolution par fusion d'images hyperspectrales et multispectrales prenant en compte la variabilité entre les images. L'image haute résolution et la variabilité inter-images sont représentées sous forme de tenseurs et sont supposées admettre une décomposition de Tucker. Deux algorithmes sont proposés, et des conditions pour une reconstruction exacte de l'image à haute résolution sont fournies. Les résultats expérimentaux illustrent les performances des méthodes proposées.

Abstract – In this paper we propose a hyperspectral and multispectral image fusion framework accounting for inter-image variability. The images are represented as three dimensional tensors, and both the high-resolution image and inter-image variations are assumed to admit a Tucker decomposition. Two algorithms are proposed, one purely algebraic and computationally fast and another based on an optimization procedure. Exact recovery conditions are provided. Experimental results illustrate the performance of the proposed methods.

1 Introduction

Combining hyperspectral images (HIs) and multispectral images (MIs) of the same scene to generate high-resolution images (HRIs) has become a popular approach to overcome physical limitations of the image acquisition process [1]. Many different approaches have been proposed to address this problem, including multiresolution analysis, which attempts to fuse high-frequency of the MI with the HI [1], and different subspace-based formulations, which consider the HRI pixels to be a linear combination of a small set of basis vectors [2, 3].

Recently, various works proposed image fusion methods that leverage the representation of the HI, MI and HRI as 3-dimensional tensors [4, 5]. This led to algorithms that allied good experimental performance with strong theoretical guarantees for the recovery of the HRI. Following this framework, several works have formulated image fusion as a coupled tensor approximation problem. In [4], a low-rank canonical polyadic decomposition was assumed for the HRI, which was recovered using an alternating optimization strategy. The algorithm was shown to recover the correct HRI when the MSI was identifiable. This was extended to consider a block-term decomposition (BTD) in order to improve the interpretability of the results in [6]. The Tucker decomposition was considered in [5], leading to low-cost closed-form algorithms based on the SVD. Exact recovery results were also obtained.

Most existing methods assume that the HI and MI are subject to the same acquisition conditions. Thus, their performance can degrade considerably when fusing images acquired by different sensors or at distinct time instants since the HI and MI may differ due to variations in, e.g., illumination and seasonal conditions [7]. This problem was recently addressed in a subspace-based formulation in [8], where the spectral basis vectors corresponding to the HI and to the MI were allowed to vary according to smooth elementwise multiplicative scalings [9]. However, the computational cost of the algorithm was high, and no theoretical guarantees were provided. Both of these issues were later addressed in [10], which used an LL1 BTD to jointly perform image fusion and unmixing of the HRI under inter-image variability. Theoretical recovery guarantees and two efficient algorithms were presented. However, these works did not consider spatially localized inter-image changes.

We propose two image fusion algorithms accounting for spatially and spectrally localized inter-image variability. Both the HRI and the inter-image variations are represented using a Tucker tensor decomposition. Two algorithms are derived, one algebraic and another based on an optimization procedure, and theoretical exact recovery results are obtained for both of them. Experiments show that the proposed methods can achieve state of the art performance in the presence of inter-image variability. Proofs of the theorems and more details can be found in the extended version of this work in [11].

Notation: We denote scalars by (x) or (X), vectors by $(\boldsymbol{x})$, matrices by $(\boldsymbol{X})$ and tensors by $(\mathcal{X})$ or (Θ). Operator $\times_k$ denotes the mode-k product between a tensor and a matrix, $[\![\mathcal{A}; \boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{B}_3]\!] = \mathcal{A} \times_1 \boldsymbol{B}_1 \times_2 \boldsymbol{B}_2 \times_3 \boldsymbol{B}_3$ the full multilinear product, and $\boldsymbol{X}_{(i)}$ the mode-i matricization of tensor $\mathcal{X}$.

2 Tensor-based imaging model

We denote an HI and an MI by $\mathcal{Y}_h \in \mathbb{R}^{N_1 \times N_2 \times L_h}$ and $\mathcal{Y}_m \in \mathbb{R}^{M_1 \times M_2 \times L_m}$, respectively, where $M_1 > N_1$ and $M_2 > N_2$ are the number of pixels in the vertical and horizontal dimensions, and $L_m < L_h$ is the number of bands. The HI and MI are considered to be spatially and spectrally degraded versions of two HRIs $\mathcal{Z}_h, \mathcal{Z}_m \in \mathbb{R}^{M_1 \times M_2 \times L_h}$, and are modeled as [4, 5]:

$$\mathcal{Y}_h = \mathcal{Z}_h \times_1 \boldsymbol{P}_1 \times_2 \boldsymbol{P}_2 + \mathcal{E}_h \,, \tag{1}$$

$$\mathcal{Y}_m = \mathcal{Z}_m \times_3 \boldsymbol{P}_3 + \mathcal{E}_m \,, \tag{2}$$

where $\mathcal{E}_m$ and $\mathcal{E}_h$ denote additive noise. Matrices $\boldsymbol{P}_1 \in \mathbb{R}^{N_1 \times M_1}$ and $\boldsymbol{P}_2 \in \mathbb{R}^{N_2 \times M_2}$ model blurring and downsampling during the HI acquisition (assumed to be separable), while $\boldsymbol{P}_3 \in \mathbb{R}^{L_m \times L_h}$ model the spectral response of the multispectral instrument [4, 5].

While most methods assume that $\mathcal{Z}_h = \mathcal{Z}_m$, HIs and MIs are often acquired by instruments on-board of different satellites, or at different dates. This introduces inter-image variability between the HI $\mathcal{Y}_h$ and the MI $\mathcal{Y}_m$ originating from, e.g., atmospheric or illumination variations, or scene changes [7]. Spatially uniform inter-image variations have been recently considered in [8]. However, more flexible models must be introduced to account for more complex types of changes. We propose to model the inter-image variability as follows:

$$\mathcal{Z}_m = \mathcal{Z}_h + \Psi \,, \tag{3}$$

where the *variability tensor* $\Psi \in \mathbb{R}^{M_1 \times M_2 \times L_h}$ represents changes occurring between the HRIs $\mathcal{Z}_h$ and $\mathcal{Z}_m$ (and thus between the HI and MI). This allows the acquisition process to be represented as:

$$\mathcal{Y}_h = \mathcal{Z}_h \times_1 \boldsymbol{P}_1 \times_2 \boldsymbol{P}_2 + \mathcal{E}_h \,, \tag{4}$$

$$\mathcal{Y}_m = (\mathcal{Z}_h + \Psi) \times_3 \boldsymbol{P}_3 + \mathcal{E}_m \,. \tag{5}$$

We are interested in recovering the HRI $\mathcal{Z}_h$ and variability tensor Ψ from the HI and MI:

$$\boxed{\text{find } \mathcal{Z}_h \in \Omega_Z,\ \Psi \in \Omega_\Psi \text{ that satisfy (4)–(5).}} \tag{6}$$

in which sets $\Omega_Z \subseteq \mathbb{R}^{M_1 \times M_2 \times L_h}$ and $\Omega_\Psi \subseteq \mathbb{R}^{M_1 \times M_2 \times L_h}$ encode prior knowledge about $\mathcal{Z}_h$ and Ψ, which is very important due to the ill-posedness of the problem. In this paper, we consider a low-rank Tucker tensor model to constrain both $\mathcal{Z}_h$ and Ψ [5]:

$$\mathcal{Z}_h = [\![\mathcal{G}_Z; \boldsymbol{B}_{Z,1}, \boldsymbol{B}_{Z,2}, \boldsymbol{B}_{Z,3}]\!] \,, \tag{7}$$

$$\Psi = [\![\mathcal{G}_\Psi; \boldsymbol{B}_{\Psi,1}, \boldsymbol{B}_{\Psi,2}, \boldsymbol{B}_{\Psi,3}]\!] \,, \tag{8}$$

in which $\boldsymbol{B}_{\pi,i} \in \mathbb{R}^{M_i \times K_{\pi,i}}$, $i \in \{1,2\}$, $\boldsymbol{B}_{\pi,3} \in \mathbb{R}^{L_h \times K_{\pi,3}}$ and $\mathcal{G}_\pi \in \mathbb{R}^{K_{\pi,1} \times K_{\pi,2} \times K_{\pi,3}}$ are factor matrices and the core tensors of the model, for $\pi \in \{Z, \Psi\}$. This reduces the amount of variables to estimate and allows recovery guarantees to be derived.

Ignoring the additive noise, combining (7)–(8) and (4)–(5) and using the properties of the mode-k product leads to [5]:

$$\mathcal{Y}_h = [\![\mathcal{G}_Z; \boldsymbol{P}_1\boldsymbol{B}_{Z,1}, \boldsymbol{P}_2\boldsymbol{B}_{Z,2}, \boldsymbol{B}_{Z,3}]\!] \,, \tag{9}$$

$$\mathcal{Y}_m = \sum_{\pi \in \{Z,\Psi\}} [\![\mathcal{G}_\pi; \boldsymbol{B}_{\pi,1}, \boldsymbol{B}_{\pi,2}, \boldsymbol{P}_3\boldsymbol{B}_{\pi,3}]\!] \tag{10}$$

$$= [\![\mathcal{C}_m; \boldsymbol{C}_{m,1}, \boldsymbol{C}_{m,2}, \boldsymbol{P}_3\boldsymbol{C}_{m,3}]\!] \,, \tag{11}$$

where $\boldsymbol{C}_{m,i} = [\boldsymbol{B}_{Z,i}\ \boldsymbol{B}_{\Psi,i}]$, $i \in \{1,2,3\}$, and $\mathcal{C}_m$ is a block-diagonal tensor with diagonal blocks $\mathcal{G}_Z$ and $\mathcal{G}_\Psi$. Note that although the MI admits a standard Tucker model, the block diagonal structure of its core tensor $\mathcal{C}_m$ plays an important role in obtaining recoverability results of the second algorithm.

3 Proposed algorithm

3.1 Algebraic solution

Since the variability tensor is only present in the MSI in (5), only a low spectral resolution $\Psi \times_3 \boldsymbol{P}_3$ can be recovered. Thus, it suffices to consider the recovery of the HRI $\mathcal{Z}_h$ from $\mathcal{Y}_h$ and $\mathcal{Y}_m$. Under model (9) and (11), for sufficiently low tensor ranks this consists in solving the following system of equations:

$$\begin{cases} \mathcal{Y}_h &= [\![\mathcal{G}_Z; \boldsymbol{C}_{h,1}, \boldsymbol{C}_{h,2}, \boldsymbol{B}_{Z,3}]\!] \\ \mathcal{Y}_m &= [\![\mathcal{C}_m; \boldsymbol{C}_{m,1}, \boldsymbol{C}_{m,2}, \boldsymbol{P}_3\boldsymbol{C}_{m,3}]\!] \\ \boldsymbol{C}_{h,i} &= \boldsymbol{P}_i\boldsymbol{B}_{Z,i}, \quad i \in \{1,2\} \\ \boldsymbol{C}_{m,i} &= \big[\boldsymbol{B}_{Z,i}, \boldsymbol{B}_{\Psi,i}\big], \quad i \in \{1,2,3\} \end{cases}, \tag{12}$$

in which $\boldsymbol{C}_{h,i}$, $i \in \{1,2\}$ are the mode-1 and mode-2 factor matrices of $\mathcal{Y}_h$. The solution $\widehat{\mathcal{Z}}_h$ is then computed using (7). An important aspect of (12) is that the block diagonal structure $\mathcal{C}_m$ is not considered. This simplifies the problem but introduces the need for more stringent assumptions to obtain a solution.

If $K_{Z,i} + K_{\Psi,i} \leqslant N_i$, $i \in \{1,2\}$, (12) can be solved algebraically using Alg. 1, named CT-STAR. It consists in estimating the HRI factor matrices using the relationship between mode-1 and mode-2 matricizations of $\mathcal{Y}_h$ and $\mathcal{Y}_m$, and recovering the core tensor separately. Theorem 1 provides exact recovery conditions for CT-STAR.

Theorem 1. *Suppose that $\mathcal{Z}_h$ and Ψ have multilinear ranks $(K_{Z,1}, K_{Z,2}, K_{Z,3})$ and $(K_{\Psi,1}, K_{\Psi,2}, K_{\Psi,3})$, respectively, that $\mathcal{Y}_h$ and $\mathcal{Y}_m$ admit Tucker decompositions as in (12), that the observation noise is zero (i.e., $\mathcal{E}_h = \mathbb{0}$, $\mathcal{E}_m = \mathbb{0}$), and that*

$$\operatorname{rank}(\boldsymbol{P}_i\boldsymbol{B}_{Z,i}) = K_{Z,i} \,, \quad i \in \{1,2\} \tag{13}$$

$$\operatorname{rank}(\boldsymbol{P}_i\boldsymbol{B}_{\Psi,i}) \leqslant K_{\Psi,i} \,, \quad i \in \{1,2\} \tag{14}$$

$$\operatorname{rank}(\boldsymbol{Y}_{h(i)}) = K_{Z,i} \,, \quad i \in \{1,2,3\} \tag{15}$$

$$\operatorname{rank}(\boldsymbol{Y}_{m(i)}) = K_{Z,i} + K_{\Psi,i} \leqslant N_i \,, \quad i \in \{1,2\} \tag{16}$$

Then, if all columns in $\boldsymbol{P}_i\boldsymbol{B}_{Z,i}$ are linearly independent from those in $\boldsymbol{P}_i\boldsymbol{B}_{\Psi,i}$, for $i \in \{1,2\}$, CT-STAR exactly recovers $\mathcal{Z}_h$.

Since CT-STAR only requires simple algebraic operations, it is computationally very efficient. However, requiring $K_{Z,i} + K_{\Psi,i} \leqslant N_i$, $i \in \{1,2\}$ limits the capability of the method to represent spatially complex HRIs, which require larger rank values. One important element which is not exploited by CT-STAR is the block-diagonal structure of the core tensor $\mathcal{C}_m$ conferred by the model (10).

Algorithm 1: CT-STAR

```
Input  : HSI $\mathcal{Y}_h$, MSI $\mathcal{Y}_m$ ranks $K_{Z,i}$, $K_{\Psi,i}$, $i \in \{1,2,3\}$
Output: HRI $\widehat{\mathcal{Z}}_h$, spectrally degraded variability factors
Check if $K_{Z,i} + K_{\Psi,i} \leqslant N_i$, $i \in \{1,2\}$ ;
$\widehat{\boldsymbol{C}}_{h,3} = \mathrm{tSVD}_{K_{Z,3}}(\boldsymbol{Y}_{h(3)})$ ;
$\widehat{\boldsymbol{C}}_{m,i} = \mathrm{tSVD}_{K_{Z,i}+K_{\Psi,i}}(\boldsymbol{Y}_{m(i)})$, for $i \in \{1,2\}$ ;
$\widetilde{\boldsymbol{Q}}_i = (\boldsymbol{P}_i\widehat{\boldsymbol{C}}_{m,i})^{\dagger}\,\mathrm{tSVD}_{K_{Z,i}}(\boldsymbol{Y}_{h(i)})$, for $i \in \{1,2\}$;
$\widetilde{\boldsymbol{C}}_{m,i} = \widehat{\boldsymbol{C}}_{m,i}\widetilde{\boldsymbol{Q}}_i$, for $i \in \{1,2\}$ ;
Compute $\widehat{\mathcal{G}}_Z$ by solving
    $(\widehat{\boldsymbol{C}}_{h,3} \otimes \boldsymbol{P}_2\widetilde{\boldsymbol{C}}_{m,2} \otimes \boldsymbol{P}_1\widetilde{\boldsymbol{C}}_{m,1})\,\mathrm{vec}(\mathcal{G}_Z) = \mathrm{vec}(\mathcal{Y}_h)$;
$\widehat{\mathcal{Z}}_h = [\![\widehat{\mathcal{G}}_Z; \widetilde{\boldsymbol{C}}_{m,1}, \widetilde{\boldsymbol{C}}_{m,2}, \widehat{\boldsymbol{C}}_{h,3}]\!]$ ;
Compute degraded variability factors as $\mathcal{Y}_m - \widehat{\mathcal{Z}}_h \times_3 \boldsymbol{P}_3$ ;
```

3.2 Optimization-based solution

Model (10) shows that when $\mathcal{E}_m = \mathbb{0}$ the MI admits a BTD [12]. This can be explored to provide recovery guarantees for the HRI under less restrictive conditions (see [11, Theorem 3] for the complete result). To this end, we formulate image fusion as the solution to the following problem:

$$\min_{\boldsymbol{\Theta}} J(\boldsymbol{\Theta}) \triangleq \left\|\mathcal{Y}_h - [\![\mathcal{G}_Z; \boldsymbol{P}_1\boldsymbol{B}_{Z,1}, \boldsymbol{P}_2\boldsymbol{B}_{Z,2}, \boldsymbol{B}_{Z,3}]\!]\right\|_F^2 + \left\|\mathcal{Y}_m - \sum_{\pi\in\{Z,\Psi\}} [\![\mathcal{G}_\pi; \boldsymbol{B}_{\pi,1}, \boldsymbol{B}_{\pi,2}, \boldsymbol{P}_3\boldsymbol{B}_{\pi,3}]\!]\right\|_F^2, \qquad (17)$$

in which $\boldsymbol{\Theta} = \{\mathcal{G}_\pi, \boldsymbol{B}_{\pi,i} : \pi \in \{Z,\Psi\}, i \in \{1,2,3\}\}$. We solve (17) using block coordinate descent (BCD), in which J is successively minimized w.r.t. $\mathcal{G}_Z$ and $\boldsymbol{B}_{Z,i}$, $i \in \{1,2,3\}$, and w.r.t. $\boldsymbol{B}_{\Psi,i}$, $i \in \{1,2,3\}$, while keeping the remaining variables fixed during each iteration. We describe this procedure in the following, and call the algorithm CB-STAR.

The optimization of J w.r.t. $\mathcal{Z}_h$ can be written as:

$$\min_{\mathcal{G}_Z, \boldsymbol{B}_{Z,i}} \left\|\mathcal{Y}_h - [\![\mathcal{G}_Z; \boldsymbol{P}_1\boldsymbol{B}_{Z,1}, \boldsymbol{P}_2\boldsymbol{B}_{Z,2}, \boldsymbol{B}_{Z,3}]\!]\right\|_F^2 + \left\|\mathcal{Y}_0 - [\![\mathcal{G}_Z; \boldsymbol{B}_{Z,1}, \boldsymbol{B}_{Z,2}, \boldsymbol{P}_3\boldsymbol{B}_{Z,3}]\!]\right\|_F^2, \qquad (18)$$

where $\mathcal{Y}_0 = \mathcal{Y}_m - [\![\mathcal{G}_\Psi; \boldsymbol{B}_{\Psi,1}, \boldsymbol{B}_{\Psi,2}, \boldsymbol{P}_3\boldsymbol{B}_{\Psi,3}]\!]$. This is a Tucker-based image fusion problem without inter-image variability, which we solve using BCD w.r.t. $\mathcal{G}_Z$ and $\boldsymbol{B}_{Z,i}$.

The optimization of J w.r.t. Ψ can be written as:

$$\min_{\mathcal{G}_\Psi, \boldsymbol{B}_{\Psi,i}, \boldsymbol{X}_2} \left\|\mathcal{Y}_1 - [\![\mathcal{G}_\Psi; \boldsymbol{B}_{\Psi,1}, \boldsymbol{B}_{\Psi,2}, \boldsymbol{X}_2]\!]\right\|_F^2, \qquad (19)$$

where $\mathcal{Y}_1 = \mathcal{Y}_m - [\![\mathcal{G}_Z; \boldsymbol{B}_{Z,1}, \boldsymbol{B}_{Z,2}, \boldsymbol{P}_3\boldsymbol{B}_{Z,3}]\!]$ and $\boldsymbol{X}_2 = \boldsymbol{P}_3\boldsymbol{B}_{\Psi,3}$. This problem is solved by computing the high-order SVD of $\mathcal{Y}_1$ [13], with rank $(K_{\Psi,1}, K_{\Psi,2}, K_{\Psi,3})$.

4 Experimental results

The proposed methods are compared to HySure [3], CNMF [2], GLPHS [14], LTMR [15], STEREO [4], SCOTT [5] and FuVar [8] using two pairs of MIs $\mathcal{Y}_m$ and HRIs $\mathcal{Z}_h$ with real inter-image variability, denominated *Lake Tahoe* and *Lockwood* (described in detail in [8]). The HRIs $\mathcal{Z}_h$ were used both to generate HIs $\mathcal{Y}_h$ according to (1) and as ground truth for the comparisons. $\boldsymbol{P}_1 = \boldsymbol{P}_2$ consisted in a Gaussian filter with unit variance and decimation by a factor of two [5], and $\mathcal{E}_h$ was a white Gaussian noise with an SNR of 30dB. $\boldsymbol{P}_3$ was obtained from calibration measurements of the Sentinel 2 instrument [8]. The reconstructed HRIs were compared to $\mathcal{Z}_h$ using four metrics: the PSNR and UIQI (for which larger values imply better reconstructions), and the SAM and ERGAS (smaller values imply better reconstructions). The metrics are defined in detail in [8]. An implementation of the proposed methods is available at `https://github.com/ricardoborsoi/CB_STAR_release`.

The results can be seen in Tables 1 and 2 and Figs. 1 and 2.

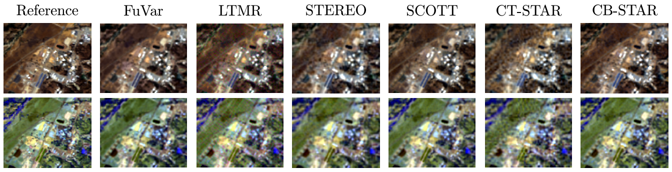

Figure 1: Visible (top) and infrared (bottom) representation of the true Lockwood HRI and its reconstructed version.

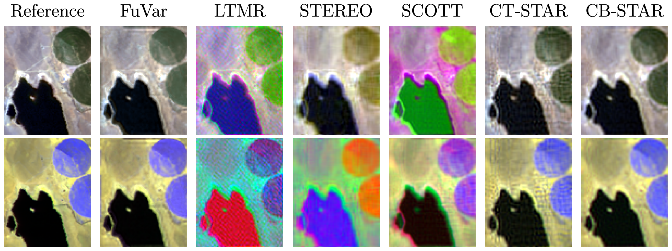

Figure 2: Visible (top) and infrared (bottom) representation of the true Lake Tahoe HRI and its reconstructed version.

Table 1: Quantitative results for the Lockwood image

Algorithm	SAM	ERGAS	PSNR	UIQI	time
HySure	3.38	7.79	23.65	0.88	4.63
CNMF	2.57	5.64	27.6	0.89	8.83
GLPHS	2.57	5.32	28.39	0.91	4.74
FuVar	2.37	**4.29**	30.59	0.95	218
LTMR	3.47	5.01	29.16	0.94	26.22
STEREO	3.49	5.51	28.72	0.93	1.14
SCOTT	2.52	4.91	29.93	0.95	**0.18**
CT-STAR	2.96	5.25	28.36	0.92	1.82
CB-STAR	**2.19**	4.35	**31.47**	**0.96**	18.8

Table 2: Quantitative results for the Lake Tahoe image

Algorithm	SAM	ERGAS	PSNR	UIQI	time
HySure	11.3	13.99	17.37	0.71	4.5
CNMF	8.79	14.59	18.37	0.71	12.1
GLPHS	5.65	7.45	24.08	0.91	4.65
FuVar	**3.91**	4.73	27.98	0.97	270.91
LTMR	34.45	1,357.42	13.8	0.52	24.94
STEREO	27.07	1,540	20.19	0.68	**0.92**
SCOTT	33.17	43,100	11.21	0.39	1.47
CT-STAR	5.41	5.25	27.25	0.96	2.88
CB-STAR	4.25	**3.78**	**30.1**	**0.98**	63.71

CB-STAR achieved the best quantitative results in all but two cases (ERGAS in Lockwood and SAM in Lake Tahoe, in which the results were similar to FuVar). Due to the limitation sin the selection of rank values, CT-STAR performed slightly worse than FuVar and CB-STAR. The methods that did not consider variability (HySure, CNMF, GLPHS, LTMR, STEREO and SCOTT) performed considerably worse, particularly in the Lake Tahoe dataset, which has a larger inter-image variability. The visual results (only shown for FuVar, LTMR, STEREO, SCOTT, CT-STAR and CB-STAR) indicate that CB-STAR achieved the closest reconstruction of the HRIs, which can be observed in the crop circles in Fig. 2, which are closer to the true HRI compared to those reconstructed by FuVar. Nevertheless, the reconstruction of the Lake Tahoe HRI by CB-STAR is smoother when compared to FuVar's. This occurs since FuVar constrains both HRIs to have the same spatial distribution of materials (which preserves sharp details from the MI), while CB-STAR allows for arbitrary spatial changes. Thus, in CB-STAR the HRI corresponding to the HI does not incorporates spatial details from the HI as effectively when large inter-image variability between the HRIs is present. CT-STAR showed some artifacts due to the limited spatial rank values, while HRIs reconstructed by the remaining methods contained significant artifacts for the Lake Tahoe dataset. The execution times of the proposed method were larger than those of SCOTT, but significantly smaller than those of FuVar.

5 Conclusions

This paper proposed an tensor-based image fusion framework accounting for inter-image variability. The HRI and the tensor representing the variability were assumed to admit a Tucker decomposition. Two algorithms were then derived, one algebraic and computationally efficient (CT-STAR) and another more flexible and based on an optimization problem (CB-STAR). Under suitable conditions, both algorithms exactly recover the HRI. Experimental results indicate that CB-STAR achieves state of the art performance in the presence of inter-image variability with lower computation times.

6 Acknowledgments

This work was supported by the CNPq, by the ANR, and by GdR ISIS.

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] 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.

[3] M. Simões, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3373–3388, 2015.

[4] 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.

[5] C. Prévost, K. Usevich, P. Comon, and D. Brie, "Hyperspectral super-resolution with coupled tucker approximation: Recoverability and SVD-based algorithms," *IEEE Transactions on Signal Processing*, vol. 68, pp. 931–946, 2020.

[6] M. Ding, X. Fu, T.-Z. Huang, J. Wang, and X.-L. Zhao, "Hyperspectral super-resolution via interpretable block-term tensor modeling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 641–656, 2020.

[7] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, C. Richard, J. Chanussot, L. Drumetz, J.-Y. Tourneret, A. Zare, and C. Jutten, "Spectral variability in hyperspectral data unmixing: A comprehensive review," *IEEE Geoscience and Remote Sensing Magazine*, pp. 2–49, 2021, doi:10.1109/MGRS.2021.3071158.

[8] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, "Super-resolution for hyperspectral and multispectral image fusion accounting for seasonal spectral variability," *IEEE Transactions on Image Processing*, vol. 29, no. 1, pp. 116–127, 2020.

[9] T. Imbiriba, R. A. Borsoi, and J. C. M. Bermudez, "Generalized linear mixing model accounting for endmember variability," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Calgary, Canada, 2018, pp. 1862–1866.

[10] C. Prévost, R. A. Borsoi, K. Usevich, D. Brie, J. C. M. Bermudez, and C. Richard, "Hyperspectral super-resolution accounting for spectral variability: Coupled tensor LL1-based recovery and blind unmixing of the unknown super-resolution image," *SIAM Journal on Imaging Sciences*, vol. 15, no. 1, pp. 110–138, 2022.

[11] R. A. Borsoi, C. Prévost, K. Usevich, D. Brie, J. C. M. Bermudez, and C. Richard, "Coupled tensor decomposition for hyperspectral and multispectral image fusion with inter-image variability," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 702–717, 2021.

[12] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms–part II: Definitions and uniqueness," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1033–1066, 2008.

[13] 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.

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

[15] R. Dian and S. Li, "Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization," *IEEE Transactions on Image Processing*, vol. 28, no. 10, pp. 5135–5146, 2019.