

BLIND HYPERSPECTRAL SUPER-RESOLUTION: COMBINING LOW RANK TENSOR AND MATRIX STRUCTURE

Charilaos I. Kanatsoulis*, Xiao Fu[†], Nicholas D. Sidiropoulos*, and Wing-Kin Ma[‡]

*Univ. of Minnesota, Minneapolis, MN, USA; [†]Oregon State Univ., Corvallis, OR, USA

^{*}Univ. of Virginia, Charlottesville, VA, USA; [‡]Chinese University of Hong Kong, Shatin, Hong Kong

ABSTRACT

Hyperspectral super-resolution refers to the task of fusing a hyperspectral image (HSI) and a multispectral image (MSI) in order to produce a super-resolution image (SRI) that has high spatial *and* spectral resolution. Popular methods leverage matrix factorization that models each pixel spectrum as a convex combination of spectral signatures belonging to few endmembers. These methods are considered state of the art, but many challenges remain. First multiband images are naturally 3-dimensional signals, while matrix methods ignore 3-d structure. Second, these methods do not provide identifiability guarantees under which the reconstruction task is feasible. Third, a tacit assumption is that the degradation operators from SRI to MSI and HSI are known or can be easily estimated – which is hardly the case in practice. Recently [1] proposed a coupled tensor factorization approach to handle these issues. In this work we propose a hybrid model that combines the benefits of tensor and matrix factorization approaches. We also develop a new algorithm that is mathematically simple, enjoys relaxed identifiability and is completely agnostic of the spatial degradation operator. Experimental results with real hyperspectral data showcase the effectiveness of the proposed approach.

Index Terms— Hyperspectral imaging, multispectral imaging, super-resolution, image fusion, tensor decomposition, identifiability

1. INTRODUCTION

Multi-sensor image fusion has long been of interest in image processing and computer vision [2, 3, 4]. In remote sensing for example, multi-band image fusion is being used to overcome physical and technical limitations of hyperspectral and multispectral sensors. In particular the task of hyperspectral super-resolution integrates information from a multispectral image (MSI), which has a fine spatial but limited spectral resolution, and a co-registered hyperspectral image (HSI), which has low spatial but high spectral resolution, to obtain a super-resolution image (SRI) that features high spatial and spectral resolution simultaneously [5, 6]. SRI's can be used to improve the performance of several analytical tasks, such as spectral unmixing, image segmentation, anomaly detection, etc. [5, 7].

The majority of existing methods approach the hyperspectral super-resolution problem from a matrix factorization perspective [8, 9, 10, 11, 12, 13, 14, 15]. Specifically the popular linear mixture model (LMM) is adopted, which asserts that each pixel spectrum of a multi-band image is a convex combination of spectral signatures of several endmembers (materials) which present at that pixel. As a result, a multi-band image can be decomposed in a bilinear form

of abundances and spectral signatures of several materials. Super-resolution is achieved by jointly decomposing a MSI and a HSI, in order to obtain a more accurate map of the materials and their abundances. Many matrix factorization based fusion methods have been proposed, using variations of the LMM, or employing different algorithmic frameworks to approach the associated non-convex optimization problem.

Although the matrix-factorization based methods are considered state of the art, several challenges remain. First and foremost is that multi-band images are inherently three dimensional (3-d) signals (2-d space $\times$ frequency) but 3-d structure is ignored in these matrix-based methods. Second, the matrix-factorization based fusion models are usually non-identifiable. Note that identifiability is important for number of reasons. It provides guarantees under which super-resolution reconstruction is feasible, as well as practitioner guidelines regarding sensing and measurement campaign design. Spectral unmixing algorithms [16, 17, 18, 19] are typically employed to initialize the fusion, and methods based on non-identifiable models are particularly vulnerable to bad initialization. Finally the vast majority of matrix factorization methods assume that the spatial degradation operation from SRI to HSI is known or can be accurately estimated, which is hardly the case in practice. While it is fair to assume that the spectral response can be obtained by inspecting the hyperspectral and multispectral sensor specifications, estimating the spatial response is challenging. The commonly adopted model of the spatial degradation operator is a blurring kernel followed by downsampling – which involves many hyperparameters, such as kernel shape and size. Existing approaches for estimating this operator, e.g., [8] require knowledge of the hyperparameters.

Recent work [1] proposed a coupled tensor factorization approach to overcome the above issues. In [1] the multi-band images are modeled as third order tensors which admit a Canonical Polyadic Decomposition (CPD). Jointly estimating the CPD factors of HSI and MSI can guarantee recovery of the SRI under mild conditions. This model takes advantage of the 3-d structure of multi-band images, is identifiable under mild conditions and can accommodate scenarios where the spatial degradation operator is completely unknown.

In this work we propose a hybrid approach that combines the benefits of the two aforementioned models. Specifically we model the super-resolution image as a low-rank tensor, while simultaneously imposing low rank matrix structure as the LMM suggests. The proposed hybrid model is identifiable and enjoys the nice properties of both models. Furthermore, we introduce a brand new hybrid algorithm, which is very appealing due to its simplicity, accuracy, and ability to work without any knowledge of the spatial degradation operator. Numerical results with real hyperspectral data show that the proposed approach outperform the state-of-the art even if they as-

Supported in part by U.S. NSF grants ECCS-1608961, IIS-1447788, and IIS-1704074.

sume knowledge of the spatial degradation.

2. PRELIMINERIES AND BACKGROUND

2.1. Tensor Algebra Preliminaries

This work uses tensor algebra, so we briefly introduce some key concepts here, and refer the reader to [20, 21] for more information. A third order tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$ is a three-way array indexed by i, j, k with elements $\underline{\mathbf{X}}(i, j, k)$. Any third order tensor can be synthesized as a sum of 3-way outer products (rank-one tensors) as in $\underline{\mathbf{X}}(i, j, k) = \sum_{f=1}^F \mathbf{A}(i, f) \mathbf{B}(j, f) \mathbf{C}(k, f)$, 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 three factors of the tensor. If F is minimum, $F = \text{rank}(\underline{\mathbf{X}})$ and the decomposition is known as canonical polyadic decomposition (CPD), or parallel factor analysis (PARAFAC). In this work, we use the notation $\underline{\mathbf{X}} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]$ to represent the CPD of the tensor.

One pivotal difference between low-rank tensor and matrix factorization is that the CPD model is essentially unique even when F is much larger than $\max\{I, J, K\}$. In particular:

Theorem 1 [22] *Let $\underline{\mathbf{X}} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]$ with $\mathbf{A} : I \times F$, $\mathbf{B} : J \times F$, and $\mathbf{C} : K \times F$. Assume that $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are drawn from a jointly continuous distribution. Let $I \geq J \geq K$ without loss of generality. If $F \leq 2^{\lfloor \log_2 J \rfloor + \lfloor \log_2 K \rfloor - 2}$, then the decomposition of $\underline{\mathbf{X}}$ in terms of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ is essentially unique, almost surely.*

Here, essential uniqueness means that if $\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}$ also satisfy $\underline{\mathbf{X}} = [\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}]$, we can only have $\mathbf{A} = \tilde{\mathbf{A}} \mathbf{\Pi} \mathbf{\Lambda}_1$, $\mathbf{B} = \tilde{\mathbf{B}} \mathbf{\Pi} \mathbf{\Lambda}_2$, and $\mathbf{C} = \tilde{\mathbf{C}} \mathbf{\Pi} \mathbf{\Lambda}_3$, where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Lambda}_i$ is a full rank diagonal matrix such that $\mathbf{\Lambda}_1 \mathbf{\Lambda}_2 \mathbf{\Lambda}_3 = \mathbf{I}$.

Two useful operations in tensor algebra are matricization and mode multiplication. There are different ways to matricize an $I \times J \times K$ tensor $\underline{\mathbf{X}}$. One is $\mathbf{X} := [\underline{\mathbf{X}}(1, 1, :), \dots, \underline{\mathbf{X}}(I, J, :)]^T \in \mathbb{R}^{IJ \times K}$, where $\underline{\mathbf{X}}(i, j, :) \in \mathbb{R}^K$ denotes a column vector containing $\underline{\mathbf{X}}(i, j, 1), \dots, \underline{\mathbf{X}}(i, j, K)$, called a third mode fiber of $\underline{\mathbf{X}}$. The CPD of a tensor can be expressed in a matrix form as $\mathbf{X} = (\mathbf{B} \odot \mathbf{A}) \mathbf{C}^T$, where $\odot$ denotes the Khatri-Rao (column-wise Kronecker) product. The mode product operator multiplies a tensor by a matrix in one mode. Mode product operations are represented as $\tilde{\mathbf{X}} = \underline{\mathbf{X}} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2 \times_3 \mathbf{P}_3$ where $\times_1$ multiplies each column of $\underline{\mathbf{X}}$ with $\mathbf{P}_1$, $\times_2$ multiplies each row of $\underline{\mathbf{X}}$ with $\mathbf{P}_2$ and $\times_3$ multiplies each fiber of $\underline{\mathbf{X}}$ with $\mathbf{P}_3$. This mode product equation can be unfolded in the form $\tilde{\mathbf{X}} = (\mathbf{P}_2 \mathbf{B} \odot \mathbf{P}_1 \mathbf{A})(\mathbf{P}_3 \mathbf{C})^T$.

2.2. Problem Statement and Background

Consider the spatially co-registered HSI cube $\underline{\mathbf{Y}}_H \in \mathbb{R}^{I_H \times J_H \times K_H}$ and MSI cube $\underline{\mathbf{Y}}_M \in \mathbb{R}^{I_M \times J_M \times K_M}$. The spatial dimensions are denoted by I_H, J_H for the HSI, by I_M, J_M for the MSI and $I_H J_H \ll I_M J_M$. The spectral dimensions are K_H, K_M for HSI and MSI, respectively, with $K_H \gg K_M$. Hyperspectral super-resolution aims at fusing a HSI and a MSI in order to produce a SRI $\underline{\mathbf{Y}}_S \in \mathbb{R}^{I_M \times J_M \times K_H}$ at the spatial dimensions of the MSI and the spectral dimension of the HSI. The remote sensing community models the degradation from SRI to MSI as a mode three multiplication with a selection and averaging matrix $\mathbf{P}_M \in \mathbb{R}^{K_M \times K_H}$ i.e.,

$$\underline{\mathbf{Y}}_M = \underline{\mathbf{Y}}_S \times_3 \mathbf{P}_M^T. \quad (1)$$

Matrix $\mathbf{P}_M$ can be estimated using the spectral specifications of the hyperspectral and multispectral sensors. As far as the degradation from SRI to HSI is concerned, the literature follows Wald's protocol [23] which models the spatial degradation of each SRI frontal slab ($\underline{\mathbf{Y}}_S(:, :, k)$) as a combination of circularly symmetric Gaussian blurring and downsampling. In the journal version of [1] we

show that this operation is separable across rows and columns of the cube and can be modeled as:

$$\underline{\mathbf{Y}}_H = \underline{\mathbf{Y}}_S \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2, \quad (2)$$

where $\mathbf{P}_1 \in \mathbb{R}^{I_H \times I_M}$, $\mathbf{P}_2 \in \mathbb{R}^{J_H \times J_M}$ and $\mathbf{P}_H = \mathbf{P}_2 \otimes \mathbf{P}_1$ is the overall spatial degradation operator ($\otimes$ denotes the Kronecker product). In practice, the blurring kernel is only approximately Gaussian, its width is only approximately known, and there is unknown sampling offset. In this work we consider $\mathbf{P}_1, \mathbf{P}_2$ to be unknown.

Matrix Factorization Approaches: The majority of existing HSI-MSI fusion methods, model multi-band images as low-rank matrices following the LMM [8, 9, 10, 11, 12, 13, 14]. The SRI, for example, can be modeled as $\mathbf{Y}_S \approx \mathbf{S} \mathbf{E}^T \in \mathbb{R}^{I_M J_M \times K_H}$, where $\mathbf{E} \in \mathbb{R}^{K_H \times R}$ is the matrix containing the spectral signatures of $R \ll \min\{I_H J_H, K_H\}$ endmembers, $\mathbf{S} \in \mathbb{R}^{I_M J_M \times R}$ is the abundance matrix, and $\mathbf{1}^T \mathbf{S}^T = \mathbf{1}^T$ and $\mathbf{S} \geq \mathbf{0}$ hold. Using the mode multiplication based forward model described by (1) and (2) we can express the matricized MSI and HSI as $\mathbf{Y}_M \approx \mathbf{S} \mathbf{E}^T \mathbf{P}_M^T \in \mathbb{R}^{I_M J_M \times K_M}$ and $\mathbf{Y}_H \approx \mathbf{P}_H \mathbf{S} \mathbf{E}^T \in \mathbb{R}^{I_H J_H \times K_H}$ respectively.

Super-resolution reconstruction is performed by estimating $\hat{\mathbf{E}}$ and $\hat{\mathbf{S}}$ via jointly factoring $\mathbf{Y}_H$ and $\mathbf{Y}_M$, and then setting $\hat{\mathbf{Y}}_S \approx \hat{\mathbf{S}} \hat{\mathbf{E}}^T$. This is the basic idea behind [8, 9, 10, 11, 12, 13, 14, 15].

Tensor Approach: Although matrix factorization-based approaches are considered state of the art, several challenges exist as previously explained. To overcome these, [1] introduced a novel coupled CPD approach to handle the hyperspectral super-resolution task. The SRI is naturally a third order tensor and admits a CPD $\underline{\mathbf{Y}}_S = [\mathbf{A}, \mathbf{B}, \mathbf{C}]$ of a certain rank F . then, following (1), (2) the MSI and HSI can be expressed as:

$$\underline{\mathbf{Y}}_M = [\mathbf{A}, \mathbf{B}, \mathbf{P}_M \mathbf{C}] \quad (3)$$

$$\underline{\mathbf{Y}}_H = [\mathbf{P}_1 \mathbf{A}, \mathbf{P}_2 \mathbf{B}, \mathbf{C}]. \quad (4)$$

The factor matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ can be estimated by joint CPD of $\underline{\mathbf{Y}}_M$ and $\underline{\mathbf{Y}}_H$ and the SRI is recovered as $\hat{\underline{\mathbf{Y}}}_S = [\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}]$. Spatially blind reconstruction can be performed by simply replacing $\mathbf{P}_1 \mathbf{A}$ and $\mathbf{P}_2 \mathbf{B}$ by auxiliary variables $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$, respectively. Compared to matrix-factorization based approaches, the CPD based super-resolution framework provides strong identifiability guarantees, and improved resolution in practice.

3. THE HYBRID MODEL

At this point, it is natural to wonder whether it is possible to come up with an approach that combines the benefits of low rank matrix factorization, as well as those of low rank tensor factorization. In this section we propose a hybrid approach that takes advantage of the multi-linear dependence across a multi-band image in conjunction with the low rank matrix structure imposed by the LMM. The SRI is a tensor that admits a CPD $\underline{\mathbf{Y}}_S = [\mathbf{A}, \mathbf{B}, \mathbf{C}]$ of rank F . Moreover, following the LMM, the mode 3 unfolding of the SRI, $\mathbf{Y}_S$, exhibits low rank matrix structure of rank R . This is reflected in the singular value decomposition (SVD) of $\mathbf{Y}_S = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$. The columns of $\mathbf{V} \in \mathbb{R}^{K_H \times R}$ are the right singular vectors of $\mathbf{Y}_S$ and give an orthogonal basis for the fiberspace of tensor $\underline{\mathbf{Y}}_S$. Then one can without loss of generality transform the original superresolution tensor $\underline{\mathbf{Y}}_S \in \mathbb{R}^{I_M \times J_M \times K_H}$ to a superresolution core tensor $\underline{\mathbf{Z}}_S \in \mathbb{R}^{I_M \times J_M \times R}$, as $\underline{\mathbf{Z}}_S = \underline{\mathbf{Y}}_S \times_3 \mathbf{V}$. The CPD model of the core tensor is:

$$\underline{\mathbf{Z}}_S = [\mathbf{A}, \mathbf{B}, \tilde{\mathbf{C}}] \quad (5)$$

where $\tilde{\mathbf{C}} = \mathbf{V}^T \mathbf{C} \in \mathbb{R}^{R \times F}$. Note that one can always recover $\underline{\mathbf{Y}}_S$ as $\underline{\mathbf{Y}}_S = \underline{\mathbf{Z}}_S \times_3 \mathbf{V}^T$ and $\mathbf{C}$ from $\mathbf{C} = \mathbf{V} \tilde{\mathbf{C}}$, since $\underline{\mathbf{Y}}_S, \mathbf{C}$

live in a low dimensional subspace defined by $\mathbf{V}$; see [20, 24] for details. The HSI is related to SRI via (2). Therefore $\underline{\mathbf{Y}}_S$, $\underline{\mathbf{Y}}_H$ share the same fiberspace $\mathbf{V}$, which can be computed by the SVD of $\mathbf{Y}_H$. As a result, one may, without loss of generality, transform the original hyperspectral tensor $\underline{\mathbf{Y}}_H$ to a hyperspectral core tensor $\underline{\mathbf{Z}}_H \in \mathbb{R}^{I_H \times J_H \times R}$ as $\underline{\mathbf{Z}}_H = \underline{\mathbf{Y}}_H \times_3 \mathbf{V} = \underline{\mathbf{Y}}_S \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2 \times_3 \mathbf{V}$. The CPD model of $\underline{\mathbf{Z}}_H$ is then:

$$\underline{\mathbf{Z}}_H = \llbracket \tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}} \rrbracket, \quad (6)$$

where $\tilde{\mathbf{A}} = \mathbf{P}_1 \mathbf{A} \in \mathbb{R}^{I_H \times F}$, $\tilde{\mathbf{B}} = \mathbf{P}_2 \mathbf{B} \in \mathbb{R}^{J_H \times F}$.

Regarding the relation between the MSI and the core SRI, $\underline{\mathbf{Y}}_M = \underline{\mathbf{Y}}_S \times_3 \mathbf{P}_M^T$ and $\underline{\mathbf{Y}}_S = \underline{\mathbf{Z}}_S \times_3 \mathbf{V}^T$. Thus, $\underline{\mathbf{Y}}_M = \underline{\mathbf{Z}}_S \times_3 \tilde{\mathbf{P}}_M^T$, where $\tilde{\mathbf{P}}_M = \mathbf{P}_M \mathbf{V} \in \mathbb{R}^{K_M \times R}$. Consequently the CPD model of the MSI can be casted as:

$$\underline{\mathbf{Y}}_M = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{P}_M \mathbf{C} \rrbracket = \llbracket \mathbf{A}, \mathbf{B}, \tilde{\mathbf{P}}_M \tilde{\mathbf{C}} \rrbracket, \quad (7)$$

Overall, the hybrid model describes the SRI and HSI by a CPD model that admits a low rank matrix structure in the third mode (fiberspace). As far as the MSI is concerned, it is also described by a CPD model and the low rank matrix structure of the third mode is reflected in the spatial degradation operator which is transformed to $\tilde{\mathbf{P}}_M = \mathbf{P}_M \mathbf{V}$. The (hybrid model based) super-resolution task is performed by identifying $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{C} = \mathbf{V} \tilde{\mathbf{C}}$.

4. SUPER-RESOLUTION CUBE ALGORITHM (SCUBA)

Taking a closer look at the hybrid model we observe that we are able to compress the spectral dimension of the HSI – and SRI since they share the same fiberspace – from K_H to R without loss of generality. The spectral response has been transformed from $\mathbf{P}_M \in \mathbb{R}^{K_M \times K_H}$ to $\tilde{\mathbf{P}}_M = \mathbf{P}_M \mathbf{V} \in \mathbb{R}^{K_M \times R}$ in the hybrid model. In practice, the number of multispectral bands is usually between $K_M = 4$ or $K_M = 8$. The number of endmembers, R , on the other hand, depends on the size and type of the image, but is usually less than 20. While this case can be successfully handled using coupled tensor factorization (as we will show in the companion journal paper), here we would like to point out a different and quite intriguing possibility. Namely, for $R \leq K_M$, super-resolution reconstruction can be accomplished in a simple and appealing way, and under relaxed identifiability conditions – even if the spatial degradation operator is non-separable and completely unknown. Let $\underline{\mathbf{Y}}_M$ denote the MSI with CPD $\underline{\mathbf{Y}}_M = \llbracket \mathbf{A}, \mathbf{B}, \tilde{\mathbf{C}} \rrbracket$, where $\tilde{\mathbf{C}} = \tilde{\mathbf{P}}_M \tilde{\mathbf{C}}$. Also let $\mathbf{V} \in \mathbb{R}^{K_H \times R}$ be the basis of the hyperspectral fiberspace computed via SVD of $\mathbf{Y}_H$. If $R \leq K_M$, $\tilde{\mathbf{C}}$ can be computed by solving the overdetermined system $\tilde{\mathbf{C}} = \tilde{\mathbf{P}}_M^\dagger \tilde{\mathbf{C}}$, and consequently $\mathbf{C}$ can be obtained as $\mathbf{C} = \mathbf{V} \tilde{\mathbf{C}}$. The procedure is summarized in the following steps:

$$\llbracket \mathbf{A}, \mathbf{B}, \tilde{\mathbf{C}} \rrbracket \leftarrow \text{CPD}(\underline{\mathbf{Y}}_M) \quad (8a)$$

$$\mathbf{V} \leftarrow \text{SVD}(\mathbf{Y}_H) \quad (8b)$$

$$\mathbf{C} = \mathbf{V} \tilde{\mathbf{P}}_M^\dagger \tilde{\mathbf{C}} \quad (8c)$$

$$\hat{\underline{\mathbf{Y}}}_S(i, j, k) = \sum_{f=1}^F \mathbf{A}(i, f) \mathbf{B}(j, f) \mathbf{C}(k, f), \quad (8d)$$

where $\dagger$ denotes Moore-Penrose pseudoinverse. The caveat is that $R \leq K_M$ is restrictive in practice. The engineering solution is to judiciously choose, e.g., $8 \times 8 \times K$ blocks of the original image tensor, similar to what is done in JPEG image compression. Small

spatial patches typically contain few endmembers, hence $R \leq K_M$ holds over each patch. Also note that smaller-size tensors typically exhibit smaller tensor rank, allowing us to use a smaller F per sub-tensor. The proposed *super-resolution cube approach* (SCUBA for short) is summarized in Algorithm 1. In the algorithm $\underline{\mathbf{Y}}_M^{(l)}$, $\underline{\mathbf{Y}}_H^{(l)}$ denote the l -th MSI, HSI cube respectively.

Algorithm 1: SCUBA

Judiciously cut $\underline{\mathbf{Y}}_M$, $\underline{\mathbf{Y}}_H$ into L cubes.
for $l=1$ to L do

$$\llbracket \mathbf{A}, \mathbf{B}, \tilde{\mathbf{C}} \rrbracket \leftarrow \text{CPD}(\underline{\mathbf{Y}}_M^{(l)})$$

$$\mathbf{V} \leftarrow \text{SVD}(\mathbf{Y}_H^{(l)})$$

$$\mathbf{C} = \mathbf{V} \tilde{\mathbf{P}}_M^\dagger \tilde{\mathbf{C}}$$

$$\hat{\underline{\mathbf{Y}}}_S^{(l)} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$$

end for

4.1. SCUBA Identifiability

While blocking may introduce artifacts in highly compressed JPEG images, this is not a concern in our context when the decomposition of each sub-tensor is identifiable – since we can then provably reconstruct each super-resolution sub-tensor independently of its neighbors.

Theorem 2 Let $\underline{\mathbf{Y}}_M = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{P}_M \mathbf{C} \rrbracket$ and $\mathbf{Y}_H = \mathbf{U} \Sigma \mathbf{V}^T$, where $R \leq K_M$. Assume without loss of generality that $I_M \geq J_M \geq K_M$. Also assume that the entries of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are drawn from some absolutely continuous distribution, and that $\mathbf{P}_H$ and $\mathbf{P}_M$ have full rank. Let $(\mathbf{A}^*, \mathbf{B}^*, \mathbf{C}^*)$ denote a solution to (8a)-(8c). Then, $\hat{\underline{\mathbf{Y}}}_S(i, j, k) = \sum_{f=1}^F \mathbf{A}^*(i, f) \mathbf{B}^*(j, f) \mathbf{C}^*(k, f)$ recovers the ground-truth $\underline{\mathbf{Y}}_S$ almost surely if $F \leq 2^{\lceil \gamma \rceil - 2}$, where $\gamma = \log_2(J_M K_M)$.

As a concrete example, consider the reconstruction of a SRI of size $128 \times 128 \times 178$ from a MSI of size $128 \times 128 \times 8$ and a HSI of size $32 \times 32 \times 178$. Theorem 2 states that reconstruction is guaranteed if the rank of the MSI satisfies $F \leq 256$. The proof of Theorem 2 uses Theorem 1 to characterize the solution of (8a). The details are relegated to the journal version.

5. SIMULATIONS

In this section we evaluate the performance of the proposed algorithm and showcase its effectiveness using real hyperspectral data. We use a publicly available HSI as if it were the SRI, and synthetically generate a HSI and a MSI in a realistic manner, following Wald's protocol [23]. In particular the degradation from SRI to HSI is modeled as the process of blurring by a 7×7 Gaussian kernel and downsampling 1 out of every $4 \times 4 = 16$ pixels of the result. The spectral response $\mathbf{P}_M$ is obtained by comparing the spectral specifications of the reference image to those of the LANDSAT [25] and QuickBird multispectral sensor, for the first and second set of experiments respectively. In order to make the model more realistic we add zero mean white Gaussian noise to the spatial and spectral degradation process. The noise variance is controlled so that the HSI has 15 dB signal to noise ratio (SNR) and the MSI 25 dB SNR. All simulations were performed in Matlab on a Linux server with 3.6GHz cores and 32GB RAM. The baseline algorithms used for comparison are: Blind STEREO [1], FUSE [10], FUMI [12], HySure [8] and CNMF [9]. Among them Blind STEREO can perform blind spatial reconstruction, assuming a separable kernel. HySure can approximately estimating the spatial response when given the kernel

size and downsampling offset. SCUBA is fully blind, even for non-separable kernels. To assess performance, we use the reconstruction SNR (R-SNR), which is defined as $10 \log \left(\frac{\|\mathbf{Y}_S\|_F^2}{\|\hat{\mathbf{Y}}_S - \mathbf{Y}_S\|_F^2} \right)$ along with cross correlation (CC), spectral angle mapper (SAM), and relative dimensional global error (ERGAS) defined in [4, 5]. In a nutshell, high values of R-SNR, CC and SAM and low values of ERGAS correspond to good super-resolution performance.

The first set of experiments uses the Cuprite HSI downloaded from the AVIRIS platform as the reference SRI. It has $\mathbf{Y}_S \in \mathbb{R}^{512 \times 614 \times 187}$. Then, $\mathbf{Y}_H \in \mathbb{R}^{128 \times 152 \times 187}$ and $\mathbf{Y}_M \in \mathbb{R}^{512 \times 614 \times 6}$ are produced. Table 1 and figure 1 show the performance of the algorithms averaged over 10 Monte Carlo simulations. The rank used for STEREO is $F = 150$ and the rank of the low rank matrix model is $R = 10$. SCUBA divides the HSI and MSI into 16 non-overlapping blocks and for each block $F = 45$ and $R = 3$. The CPD is performed by Tensorlab [26], which runs for at most 25 iterations. The second set of experiments uses a hyperspectral scene taken from Pavia University in Italy captured by the ROSIS sensor. The SRI is $\mathbf{Y}_S \in \mathbb{R}^{608 \times 336 \times 103} \rightarrow \mathbf{Y}_H \in \mathbb{R}^{152 \times 84 \times 103}$, $\mathbf{Y}_M \in \mathbb{R}^{608 \times 336 \times 4}$. For STEREO we use $F = 300$ and $R = 9$; for SCUBA we cut the images into 16 pieces and use $F = 120$ and $R = 3$ for each.

Summarizing the results, SCUBA shows the best super-resolution performance, whereas the previously proposed STEREO comes second. The results are even more remarkable if one notes that SCUBA and STEREO work without knowing the spatial degradation, while HySure is given the kernel size and downsampling offset, and FUSE, FUMI and CNMF assume perfect knowledge of the spatial degradation. As far as time is concerned, FUSE is the fastest but gives low quality results. SCUBA and STEREO are the second fastest and SCUBA can even be fully parallelized across sub-tensor blocks.

Table 1: Performance of the algorithms in Cuprite Data.

Algorithm	R-SNR	CC	SAM	ERGAS	runtime (sec)
STEREO	27.88	0.9381	1.8004	1.1044	13.5
SCUBA	29.06	0.9521	1.4695	0.9714	15
FUSE	18.14	0.6952	6.4971	3.4517	1.5
FUMI	24.04	0.8707	3.1939	3.1939	259
HySure	24.20	0.8808	3.1095	1.6816	148
CNMF	22.97	0.8590	3.6957	1.9614	71.5

Table 2: Performance of the algorithms in Pavia University data

Algorithm	NMSE	CC	SAM	ERGAS	runtime (sec)
STEREO	20.39	0.9732	5.8279	3.3333	28
SCUBA	22.84	0.9843	4.31	2.5624	20.5
FUSE	19.43	0.96648	6.9954	3.9471	0.6
FUMI	22.01	0.9811	4.37	2.7328	116
HySure	19.89	0.9723	5.7094	3.1862	85
CNMF	18.43	0.9656	6.3049	3.9316	20

6. CONCLUSIONS

We proposed a novel hybrid super-resolution approach that combines the benefits of low-rank matrix and tensor approaches. We also developed a simple new algorithm named SCUBA, based on tensor block partitioning. SCUBA is computationally appealing, trivial to parallelize across blocks, and fully blind in terms of the (possibly non-separable) spatial degradation – while retaining strong identifiability properties. Relative to STEREO, SCUBA further requires knowledge of (a bound on) the number of endmembers in each block, so the two are complementary to each other, in this sense.

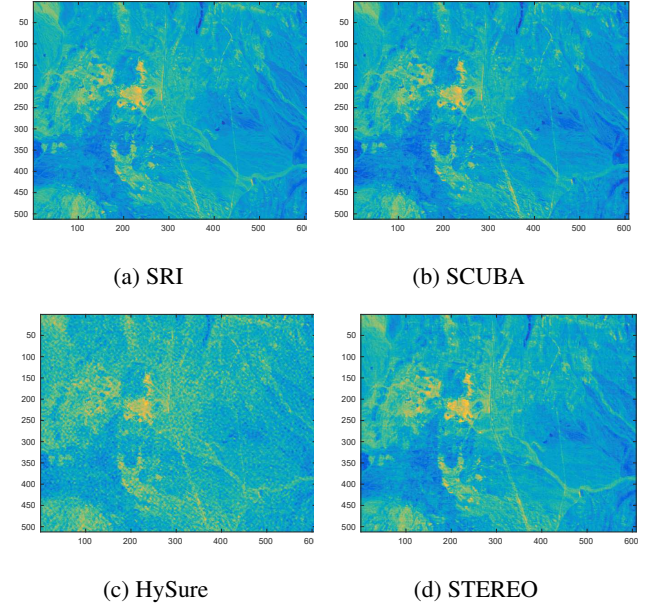


Fig. 1: Cuprite Reconstruction

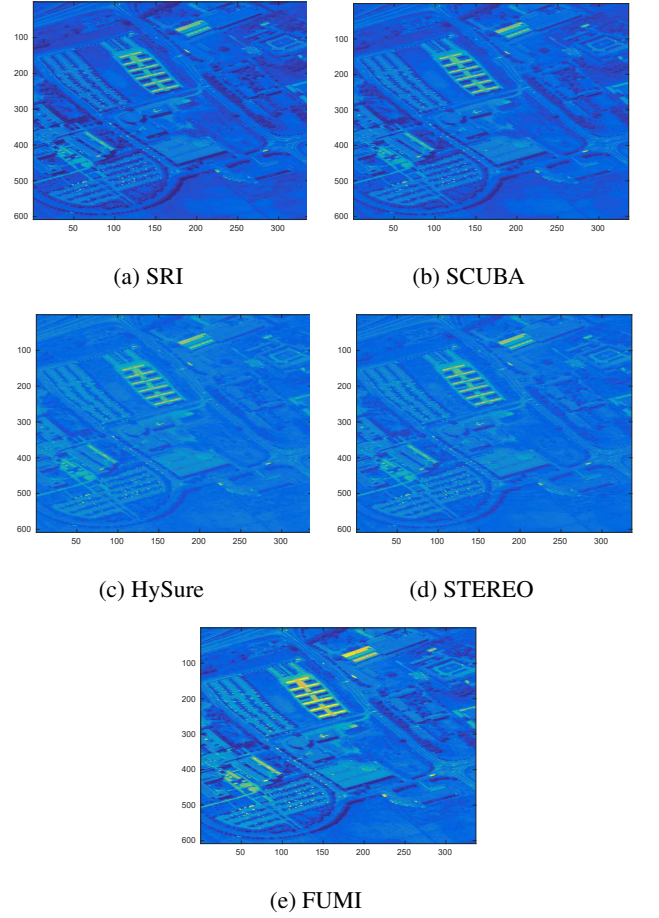


Fig. 2: Pavia University Reconstruction

7. REFERENCES

- [1] C. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Hyperspectral super-resolution via coupled tensor factorization: Identifiability and algorithms," in *Proc. IEEE ICASSP 2018*, Calgary, Canada, Apr. 15-20.
- [2] H. Li, B. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical models and image processing*, vol. 57, no. 3, pp. 235–245, 1995.
- [3] C. Pohl and J. L. Van Genderen, "Review article multisensor image fusion in remote sensing: concepts, methods and applications," *International journal of remote sensing*, vol. 19, no. 5, pp. 823–854, 1998.
- [4] L. Wald, *Data fusion: definitions and architectures: fusion of images of different spatial resolutions*. Presses des MINES, 2002.
- [5] 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.
- [6] 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 *et al.*, "Hyperspectral pansharpening: A review," *IEEE Geoscience and remote sensing magazine*, vol. 3, no. 3, pp. 27–46, 2015.
- [7] G. A. Shaw and H. K. Burke, "Spectral imaging for remote sensing," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 3–28, 2003.
- [8] 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.
- [9] 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.
- [10] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Fast fusion of multi-band images based on solving a sylvester equation," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4109–4121, 2015.
- [11] 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.
- [12] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, M. Chen, and S. Godsill, "Multiband image fusion based on spectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7236–7249, 2016.
- [13] M. A. Veganzones, M. Simoes, G. Licciardi, N. Yokoya, J. M. Bioucas-Dias, and J. Chanussot, "Hyperspectral super-resolution of locally low rank images from complementary multisource data," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 274–288, 2016.
- [14] E. Wycoff, T.-H. Chan, K. Jia, W.-K. Ma, and Y. Ma, "A non-negative sparse promoting algorithm for high resolution hyperspectral imaging," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1409–1413.
- [15] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.
- [16] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. 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, no. 2, pp. 354–379, April 2012.
- [17] —, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 354–379, 2012.
- [18] W. K. Ma, J. M. Bioucas-Dias, T. H. Chan, N. Gillis, P. Gader, A. J. Plaza, A. Ambikapathi, and C. Y. Chi, "A signal processing perspective on hyperspectral unmixing: Insights from remote sensing," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 67–81, Jan 2014.
- [19] T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4177–4193, 2011.
- [20] 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.
- [21] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [22] L. Chiantini and G. Ottaviani, "On generic identifiability of 3-tensors of small rank," *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 3, pp. 1018–1037, 2012.
- [23] 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, pp. 691–699, 1997.
- [24] I. Domanov and L. D. Lathauwer, "Canonical polyadic decomposition of third-order tensors: Reduction to generalized eigenvalue decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 35, no. 2, pp. 636–660, 2014.
- [25] B. L. Markham, J. C. Storey, D. L. Williams, and J. R. Irons, "Landsat sensor performance: history and current status," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 12, pp. 2691–2694, 2004.
- [26] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab v3. 0, march 2016," URL: <http://www.tensorlab.net>.