Robust Compressive Spectral Image Recovery Algorithm Using Dictionary Learning and Transform Tensor SVD

Yesid Fonseca
dept. of mathematics
Universidad Industrial de Santander
Bucaramanga, Colombia
yesid.fonseca@correo.uis.edu.co

Tatiana Gelvez
dept. of electrical engineering
Universidad Industrial de Santander
Bucaramanga, Colombia
tatiana.gelvez@correo.uis.edu.co

Henry Arguello Fuentes
dept. of computer science
Universidad Industrial de Santander
Bucaramanga, Colombia
henarfu@uis.edu.co

***Abstract*—This paper proposes a low-rank tensor minimization algorithm to recover a spectral image (SI) from a set of compressed observations. The proposal takes advantage of the transform tensor singular value decomposition (tt-SVD) to promote a low-rank structure on the recovered SI. The methodology has three stages. First, a poor low-rank version of the SI is estimated using the tt-SVD framework with the discrete cosine transform (DCT). Then, an orthogonal transform is learned from the initial estimation using dictionary learning. Finally, an algorithm to find a low-rank approximation of the SI in both, the DCT and the learned transform is introduced. Quantitative evaluation over two databases and two compressive optical systems shows that the proposed method improves the reconstruction quality in up to 10dB as well as it is robust in the presence of noise.**

***Index Terms*—Compressive spectral imaging, Transform tensor singular value decomposition, Dictionary learning.**

I. Introduction

A spectral image (SI) collects the intensity of reflected light at various spatial-spectral positions of a scene, which is useful for detection and classification tasks [1]. Compressive spectral imaging (CSI) ably acquires the most significant information of a SI with few random compressed projections [2]. Mathematically, let $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}$ be a third-order tensor that represents a SI with $N_1 \times N_2$ spatial pixels and N_3 spectral bands. Further, let $\mathcal{A} : \mathbb{C}^{N_1 \times N_2 \times N_3} \to \mathbb{R}^m$ be a linear operator that models a CSI architecture that obtains $\mathbf{y} \in \mathbb{C}^m$ a set of $m \ll N_1 N_2 N_3$ compressed measurements. Thus, the CSI process can be written as $\mathbf{y} = \mathcal{A}(\boldsymbol{\mathcal{F}}) + \boldsymbol{\omega}$, where $\boldsymbol{\omega} \in \mathbb{C}^m$ denotes additive Gaussian noise.

Recovering the SI from the observations $\mathbf{y} \in \mathbb{C}^m$ is a very ill-posed problem and so forth realistic priors should be included to find a feasible solution [3]. Low-rank is a prior related to the high spatial-spectral correlations that are preserved when modelling a SI as a tensor [4]. This prior suggests that a SI can be represented with few linearly independent atoms [5]. Then, it can be used to solve tensor inverse problems such as tensor completion by promoting a low-rank structure [6]. Nonetheless, the concept of rank in a tensor varies according to the definition [7]. For instance, the definition of the matrix trace can be extended to the tensor trace for getting a metric of the rank of the tensor [8], the canonical polyadic or CANDECOMP/PARAFAC method, which decomposes a tensor as the sum of rank-one outer products [9]; the tucker decomposition expresses a tensor as the product of small matrices and a tensor core [10]; the tensor singular value decomposition (t-SVD) decomposes a tensor as the t-product of three tensors using the Fourier transform [11]; and the transform tensor singular value decomposition (tt-SVD) generalizes the t-SVD decomposition by replacing the Fourier transform for any orthogonal transform [6].

This paper proposes a methodology of three stages which takes advantage of the tt-SVD to recover a SI from a set of compressed observations. The tt-SVD is employed since it preserves the 3D data structure and it allows us to find a domain in which the tensor exhibits a lower-rank, by choosing an implicit orthogonal transform. Specifically, the first stage estimates a poor low-rank version of the SI from the compressed observations by using the tt-SVD with the discrete cosine transform (DCT). Particularly, the DCT is employed since it shows high performance when condensing the most relevant information of a SI [12], [13]. The second stage learns an orthogonal transform to better represent the SI by using an orthogonal dictionary learning algorithm [14]. The third stage formulates a CSI recovery inverse problem which jointly induces a low-rank representation in both, the DC, and the learned orthogonal transform domain. Experimental results show that the reconstruction quality improves using only the learned transform in the tt-SVD framework but just in the noiseless case. On the other hand, the reconstruction quality shows a stable behaviour when only the DCT transform is used for noisy measurements. Therefore, the main contribution of the paper is a reconstruction algorithm that improves the image quality and is robust in the presence of noise. Quantitative simulation results over two datasets and two CSI optical systems show that promoting a low-rank tensor structure on both, the DCT and the learned transform domains, improves the quality of the CSI recoveries in up to 10dB in terms of spatial peak signal to noise ratio.

978-9-0827-9703-9/19/$31.00

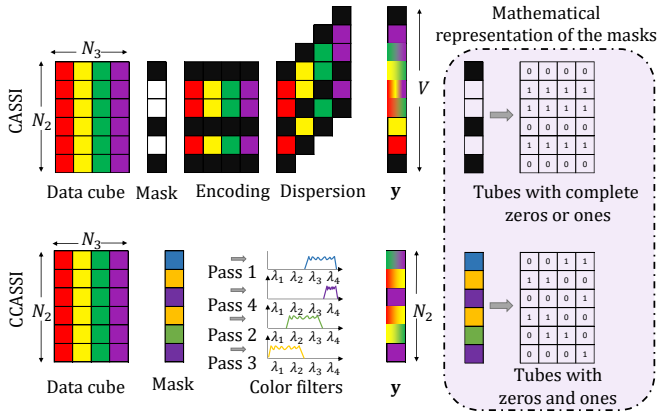

Fig. 1. CSI remarkable architectures. The CASSI mask uses opaque and translucent pieces codifying the spatial dimension. The CCASSI mask uses color filters codifying the spatial-spectral dimension.

II. Compressive Spectral Imaging Optical Systems

The coded aperture snapshot spectral imager (CASSI) [1], and the colored coded aperture snapshot spectral imager (CCASSI) [15], [16], are two CSI architectures that acquire sets of measurements in a single snapshot using a mask to encode the information. Figure 1 shows a top view of the physical acquisition process for a single spatial-spectral slice using CASSI, and CCASSI, to acquire a SI with $N_1 \times N_2 = 6 \times 6$ spatial pixels, and $N_3 = 4$ spectral bands. In CASSI, the slice is spatially encoded and spectrally dispersed before being integrated in the sensor, where $m_{CASSI} = N_1(N_2 + N_3 - 1)$ measurements. In CCASSI, the slice is spatially and spectrally encoded with a mask that contains color optical filters. Then, the light is integrated into the sensor obtaining $m_{CCASSSI} = N_1 N_2$ measurements. Multiple snapshots can be obtained by using a different masks in both architectures.

III. Tensor Rank Definition

This section defines the rank of a tensor using an orthogonal transform $\Theta : \mathbb{C}^{N_3} \to \mathbb{C}^{N_3}$ as presented in [6], [17].

Definition 1. *Θ-transform of a tensor*. Given the third-order tensor $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}$, the Θ-transform of $\boldsymbol{\mathcal{F}}$ is the tensor $\hat{\boldsymbol{\mathcal{F}}}_\Theta \in \mathbb{C}^{N_1 \times N_2 \times N_3}$ defined as

$$\hat{\boldsymbol{\mathcal{F}}}_\Theta(n_1, n_2, :) = \Theta(\boldsymbol{\mathcal{F}}(n_1, n_2, :)), \tag{1}$$

where $\boldsymbol{\mathcal{F}}(n_1, n_2, :)$ denotes the tube $\mathbf{f}^{(n_1,n_2)} = [\boldsymbol{\mathcal{F}}(n_1, n_2, 1) \ \boldsymbol{\mathcal{F}}(n_1, n_2, 2) \ \ldots \ \boldsymbol{\mathcal{F}}(n_1, n_2, N_3)]^T$ for $n_1 = 1, \ldots, N_1$, and $n_2 = 1, \ldots, N_2$. This is a linear and invertible operation.

Definition 2. *Block diagonal matrix for a tensor*. It is denoted by $\bar{\mathbf{F}}_\Theta \in \mathbb{C}^{N_1 N_3 \times N_2 N_3}$, and can be expressed as

$$\bar{\mathbf{F}}_\Theta = diag(\bar{\mathbf{F}}_\Theta^{(1)}, \bar{\mathbf{F}}_\Theta^{(2)}, \cdots, \bar{\mathbf{F}}_\Theta^{(N_3)}), \tag{2}$$

where $\bar{\mathbf{F}}_\Theta^{(n_3)}$ is a matrix in $\mathbb{C}^{N_1 \times N_2}$ corresponding to the n_3-th frontal slice of $\hat{\boldsymbol{\mathcal{F}}}_\Theta$, i.e., its entries are given by $\bar{\mathbf{F}}_\Theta^{(n_3)}(n_1, n_2) = \hat{\boldsymbol{\mathcal{F}}}_\Theta(n_1, n_2, n_3)$, for $n_3 = 1, \ldots, N_3$.

Definition 3. *Tensor rank*. Let $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}$ be a tensor. The tensor rank (t-rank) denoted by $r :=$ t-rank$(\boldsymbol{\mathcal{F}})$ is defined as the maximum rank of each slice of $\hat{\mathcal{F}}_\Theta$, i.e.,

$$r := \text{t-rank}(\boldsymbol{\mathcal{F}}) = \max_{1 \le n_3 \le N_3} \text{rank}(\hat{\boldsymbol{\mathcal{F}}}_\Theta(:, :, n_3)). \tag{3}$$

Definition 4. Given a tensor $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}$ its Nuclear norm is given by

$$\|\boldsymbol{\mathcal{F}}\|_* := \left\|\bar{\mathbf{F}}_\Theta\right\|_* = \sum_{1 \le n_3 \le N_3} \left\|\hat{\boldsymbol{\mathcal{F}}}_\Theta(:, :, n_3)\right\|_*. \tag{4}$$

IV. Robust CSI Recovery Algorithm Based on TT-SVD and Dictionary Learning

The proposed method to recover a SI is based on tt-SVD to jointly promote a low-rank structure over the DCT and a learned transform domain, consists of three main stages. The first stage encompasses an initial estimation of the SI $\hat{\boldsymbol{\mathcal{F}}}^0 \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ from the compressed measurements $\mathbf{y} \in \mathbb{C}^m$ by using the tt-SVD approach with the DCT, denoted by $\Theta_1 : \mathbb{C}^{N_3} \to \mathbb{C}^{N_3}$. The second stage involves the learning of an orthogonal transform, denoted by $\Theta_2 : \mathbb{C}^{N_3} \to \mathbb{C}^{N_3}$, using the initial estimation $\hat{\boldsymbol{\mathcal{F}}}^0$. The third stage provides the recovery of the SI by solving the proposed inverse problem, which looks for a low-rank representation of the SI in both, the DCT Θ_1 and the learned transform Θ_2 domains.

A. Stage 1: Initial Estimation

Given the set of compressed observations $\mathbf{y} = \mathcal{A}(\boldsymbol{\mathcal{F}}) \in \mathbb{C}^m$ of the SI $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}$, acquired with the compressive linear operator $\mathcal{A} : \mathbb{C}^{N_1 \times N_2 \times N_3} \to \mathbb{C}^m$, and the DCT, $\Theta_1 : \mathbb{C}^{N_3} \to \mathbb{C}^{N_3}$. The CSI inverse problem based on tt-SVD to find the initial estimation $\hat{\boldsymbol{\mathcal{F}}}^0$ aims at finding a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}$, such that, t-rank$(\boldsymbol{\mathcal{X}}) \le$ t-rank$(\boldsymbol{\mathcal{F}})$, as well as, $\mathcal{A}(\boldsymbol{\mathcal{X}}) = \mathbf{y}$, by solving

$$\begin{aligned} \underset{\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}}{\text{minimize}} \quad & \text{t-rank}(\boldsymbol{\mathcal{X}}) \\ \text{subject to} \quad & \mathcal{A}(\boldsymbol{\mathcal{X}}) = \mathbf{y}. \end{aligned} \tag{5}$$

To solve the problem in (5) which is not convex, let us use the fact that authors in [6] proved that the tensor nuclear norm $\|\cdot\|_*$ is the convex envelope of the tensor rank, t-rank$(\cdot)$. In consequence, the nearest convex problem to (5) can be written as

$$\begin{aligned} \underset{\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}}{\text{minimize}} \quad & \|\boldsymbol{\mathcal{X}}\|_* \\ \text{subject to} \quad & \mathcal{A}(\boldsymbol{\mathcal{X}}) = \mathbf{y}. \end{aligned} \tag{6}$$

To write (6) in a convenient form, let Ω be the set of all block-diagonal matrices which have N_3 matrices of dimension $N_1 \times N_2$ in their diagonal, i.e., if $\bar{\mathbf{X}} \in \Omega$, then $\bar{\mathbf{X}}$ has the following form

$$\bar{\mathbf{X}} = diag(\bar{\mathbf{X}}^{(1)}, \bar{\mathbf{X}}^{(2)}, \cdots, \bar{\mathbf{X}}^{(N_3)}), \tag{7}$$

where $\bar{\mathbf{X}}^{(k)} \in \mathbb{C}^{N_1 \times N_2}$ for $k = 1, \ldots, N_3$. It is clear that the space of tensor $\mathbb{C}^{N_1 \times N_2 \times N_3}$ is isomorphic to Ω, because $\dim(\mathbb{C}^{N_1 \times N_2 \times N_3}) = N_1 N_2 N_3 = \dim(\Omega)$.

Let $\mathcal{G}_{\Theta_1} : \mathbb{C}^{N_1 \times N_2 \times N_3} \to \Omega$ denote the linear injective operator that converts the third-order tensor in its block diagonal form with respect to the transform Θ_1, i.e., $\mathcal{G}_{\Theta_1}(\boldsymbol{\mathcal{F}}) = \bar{\mathbf{F}}_{\Theta_1}$. Then, by definition 4, we have that $\|\boldsymbol{\mathcal{X}}\|_* = \|\mathcal{G}_{\Theta_1}(\boldsymbol{\mathcal{X}})\|_*$, thus, problem in (6) is equivalent to

$$\begin{aligned} \underset{\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}}{\text{minimize}} \quad & \|\mathcal{G}_{\Theta_1}(\boldsymbol{\mathcal{X}})\|_* \\ \text{subject to} \quad & \mathcal{A}(\boldsymbol{\mathcal{X}}) = \mathbf{y}, \end{aligned} \tag{8}$$

which is also equivalent to the following convenient problem

$$\begin{aligned} \underset{\bar{\mathbf{X}} \in \Omega}{\text{minimize}} \quad & \left\|\bar{\mathbf{X}}\right\|_* \\ \text{subject to} \quad & \mathcal{A} \circ \mathcal{G}_{\Theta_1}^{-1}(\bar{\mathbf{X}}) = \mathbf{y}. \end{aligned} \tag{9}$$

Problem in (9) is efficiently solved in [18].

B. Stage 2: Learning of a Transform Operator

As the kernel of the tensor rank framework presented in section III is an orthogonal transform, the learning of a transform operator to represent the SI was carried out through the orthogonal dictionary learning technique presented in [14]. This methodology designs a dictionary that better fits a sparsity model using a learning process. In particular, we employ the initial estimation of the SI $\hat{\boldsymbol{\mathcal{F}}}^0$ as the input to learn the orthogonal transform $\Theta_2 : \mathbb{C}^{N_3} \to \mathbb{C}^{N_3}$.

C. Stage 3: CSI Inverse Problem for Final Estimation

This paper proposes an inverse problem which looks for a low-rank approximation of the SI in the DCT Θ_1 and the learned orthogonal transform Θ_2 domains. For this, problem in (8) can be rewritten as

$$\begin{aligned} \underset{\bar{\mathbf{X}} \in \Omega,\ \boldsymbol{\mathcal{W}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}}{\text{minimize}} \quad & \left\|\bar{\mathbf{X}}\right\|_* \\ \text{subject to} \quad & \bar{\mathbf{X}} = \mathcal{G}_{\Theta_1}(\boldsymbol{\mathcal{W}}),\ \ \mathcal{A}(\boldsymbol{\mathcal{W}}) = \mathbf{y}. \end{aligned} \tag{10}$$

Therefore, the proposed inverse problem in this paper consists on the modification of problem (10) as

$$\begin{aligned} \underset{\substack{\bar{\mathbf{X}}, \bar{\mathbf{Z}} \in \Omega, \\ \boldsymbol{\mathcal{W}} \in \mathbb{C}^{N_1 \times N_2 \times N_3}}}{\text{minimize}} \quad & \left\|\bar{\mathbf{X}}\right\|_* + \lambda \left\|\bar{\mathbf{Z}}\right\|_* \\ \text{subject to} \quad & \bar{\mathbf{X}} = \mathcal{G}_{\Theta_1}(\boldsymbol{\mathcal{W}}), \quad \bar{\mathbf{Z}} = \mathcal{G}_{\Theta_2}(\boldsymbol{\mathcal{W}}), \\ & \mathcal{A}(\boldsymbol{\mathcal{W}}) = \mathbf{y}. \end{aligned} \tag{11}$$

Notice that, the augmented Lagrangian function of (11) is

$$\begin{aligned} \mathcal{L}_\beta(\bar{\mathbf{X}}, \bar{\mathbf{Z}}, \boldsymbol{\mathcal{W}}, \bar{\mathbf{D}}_1, \bar{\mathbf{D}}_2, \mathbf{d}) = & \left\|\bar{\mathbf{X}}\right\|_* + \lambda \left\|\bar{\mathbf{Z}}\right\|_* \\ & + \frac{\beta}{2} \left\|\bar{\mathbf{X}} - \mathcal{G}_{\Theta_1}(\boldsymbol{\mathcal{W}}) + \bar{\mathbf{D}}_1\right\|_F^2 \\ & + \frac{\beta}{2} \left\|\bar{\mathbf{Z}} - \mathcal{G}_{\Theta_2}(\boldsymbol{\mathcal{W}}) + \bar{\mathbf{D}}_2\right\|_F^2 \\ & + \frac{\beta}{2} \left\|\mathcal{A}(\boldsymbol{\mathcal{W}}) - \mathbf{y} + \mathbf{d}\right\|_F^2, \end{aligned} \tag{12}$$

where $\bar{\mathbf{D}}_1, \bar{\mathbf{D}}_2 \in \Omega$ and $\mathbf{d} \in \mathbb{C}^m$ are the scaled dual variables, and $\beta > 0$ is the penalty parameter. In this manner, to solve problem (11) we use the following extended alternating direction method of multipliers (ADMM) [19] iterative scheme

$$\bar{\mathbf{X}}^{k+1} = \operatorname{argmin} \mathcal{L}_\beta(\bar{\mathbf{X}}, \bar{\mathbf{Z}}^k, \boldsymbol{\mathcal{W}}^k, \bar{\mathbf{D}}_1^k, \bar{\mathbf{D}}_2^k, \mathbf{d}^k), \tag{13}$$

$$\bar{\mathbf{Z}}^{k+1} = \operatorname{argmin} \mathcal{L}_\beta(\bar{\mathbf{X}}^{k+1}, \bar{\mathbf{Z}}, \boldsymbol{\mathcal{W}}^k, \bar{\mathbf{D}}_1^k, \bar{\mathbf{D}}_2^k, \mathbf{d}^k), \tag{14}$$

$$\boldsymbol{\mathcal{W}}^{k+1} = \operatorname{argmin} \mathcal{L}_\beta(\bar{\mathbf{X}}^{k+1}, \bar{\mathbf{Z}}^{k+1}, \boldsymbol{\mathcal{W}}, \bar{\mathbf{D}}_1^k, \bar{\mathbf{D}}_2^k, \mathbf{d}^k), \tag{15}$$

$$\bar{\mathbf{D}}_1^{k+1} = \bar{\mathbf{D}}_1^k + \bar{\mathbf{X}}^{k+1} - \mathcal{G}_{\Theta_1}(\boldsymbol{\mathcal{W}}^{k+1}), \tag{16}$$

$$\bar{\mathbf{D}}_2^{k+1} = \bar{\mathbf{D}}_2^k + \bar{\mathbf{Z}}^{k+1} - \mathcal{G}_{\Theta_2}(\boldsymbol{\mathcal{W}}^{k+1}), \tag{17}$$

$$\mathbf{d}^{k+1} = \mathbf{d}^k + \mathcal{A}(\boldsymbol{\mathcal{W}}^{k+1}) - \mathbf{y}. \tag{18}$$

Observe that, convergence of the previous extended ADMM is ensured by the orthogonality condition of the linear constraint in (11) [20]. Further, subproblem (15) is a Frobenius norm minimization problem with closed form solution. Finally, subproblems (13) and (14) have been efficiently solved in [18] exploiting the following concept. Let $\mathcal{D}_\delta : \Omega \to \Omega$ be the *shrinkage* operator which makes a soft threshold to the singular values, i.e., given $\bar{\mathbf{Y}} \in \Omega$ with SVD $\bar{\mathbf{Y}} = \bar{\mathbf{U}}\bar{\mathbf{S}}\bar{\mathbf{V}}^*$, with $\bar{\mathbf{U}}, \bar{\mathbf{S}}, \bar{\mathbf{V}} \in \Omega$, then $\mathcal{D}_\delta(\bar{\mathbf{Y}}) := \bar{\mathbf{U}}(\bar{\mathbf{S}} - \delta I)_+ \bar{\mathbf{V}}^*$, where $(a)_+ := \max\{0, a\}$. In [21], [22] it was proved that the shrinkage operator provides the solution of the problem

$$\mathcal{D}_\delta(\bar{\mathbf{Y}}) = \underset{\bar{\mathbf{X}} \in \Omega}{\operatorname{argmin}}\ \delta \left\|\bar{\mathbf{X}}\right\|_* + \frac{1}{2} \left\|\bar{\mathbf{X}} - \bar{\mathbf{Y}}\right\|_F^2. \tag{19}$$

In this manner Algorithm 1 summarizes the extended ADDMM iterative scheme.

Algorithm 1 Robust CSI recovery based on tt-SVD

procedure CSI–SVD($\mathcal{A} : \mathbb{C}^{N_1 \times N_2 \times N_3} \to \mathbb{C}^m$; $\mathbf{y} \in \mathbb{C}^m$; $\lambda, \beta > 0$; n; $\Theta_1, \Theta_2 : \mathbb{C}^{N_3} \to \mathbb{C}^{N_3}$)
  $k \leftarrow 0$, $\boldsymbol{\mathcal{W}}^0 \leftarrow 0$, $\bar{\mathbf{D}}_1{}^0 \leftarrow 0$, $\bar{\mathbf{D}}_2{}^0 \leftarrow 0$, $\mathbf{d}^0 \leftarrow 0$
  while $k < n$ **do**
    $\bar{\mathbf{X}}^{k+1} \leftarrow \mathcal{D}_{1/\beta}(\mathcal{G}_{\Theta_1}(\boldsymbol{\mathcal{W}}^k) - \bar{\mathbf{D}}_1^k)$.
    $\bar{\mathbf{Z}}^{k+1} \leftarrow \mathcal{D}_{\lambda/\beta}(\mathcal{G}_{\Theta_2}(\boldsymbol{\mathcal{W}}^k) - \bar{\mathbf{D}}_2^k)$.
    $\boldsymbol{\mathcal{W}}^{k+1} \leftarrow \operatorname{argmin} \mathcal{L}_\beta(\bar{\mathbf{X}}^{k+1}, \bar{\mathbf{Z}}^{k+1}, \boldsymbol{\mathcal{W}}, \bar{\mathbf{D}}_1^k, \bar{\mathbf{D}}_2^k, \mathbf{d}^k)$
    $\bar{\mathbf{D}}_1^{k+1} \leftarrow \bar{\mathbf{D}}_1^k + \bar{\mathbf{X}}^{k+1} - \mathcal{G}_{\Theta_1}(\boldsymbol{\mathcal{W}}^{k+1})$.
    $\bar{\mathbf{D}}_2^{k+1} \leftarrow \bar{\mathbf{D}}_2^k + \bar{\mathbf{Z}}^{k+1} - \mathcal{G}_{\Theta_2}(\boldsymbol{\mathcal{W}}^{k+1})$.
    $\mathbf{d}^{k+1} \leftarrow \mathbf{d}^k + \mathcal{A}(\boldsymbol{\mathcal{W}}^{k+1}) - \mathbf{y}$.
  end while
  return $\boldsymbol{\mathcal{W}}^n$
end procedure

V. Simulations and Results

A. Experimental Setup

The compressed measurements were obtained simulating the CASSI and CCASSI systems. Experiments vary the amount of captured data which refers to the ratio between the observations with respect to the total amount of pixels, i.e.,

$$\%\ \text{of data} = \frac{m}{N_1 N_2 N_3}. \tag{20}$$

Further, experiments vary the noise level which was included in the measurements as additive Gaussian noise. This noise was quantified in terms of the signal to noise ratio (SNR),

Fig. 2. RGB map of toys, Urban, and Jasper HSI databases.

measured in decibels [dB], which compares the level of the signal to the level of noise, and is given by

$$\text{SNR} = \frac{10\log_{10}(P_{\text{signal}})}{10\log_{10}(P_{\text{noise}})}, \tag{21}$$

where, P denotes average power.

The experiments were carried out over two databases: a scene containing four toys, $\mathcal{F}_1 \in \mathbb{R}^{149\times199\times32}$ captured at our laboratory; and a section of the Jasper database, $\mathcal{F}_2 \in \mathbb{R}^{100\times100\times50}$ taken from [23]. An RGB visualization of these databases is shown in Fig. 2

Qualitative evaluation is based on the peak signal to noise ratio (PSNR). The spatial PSNR of the estimated SI is calculated as the average between the PSNR_{n_3} for $n_3 = 1, \ldots, N_3$ at each spectral band. The PSNR for the n_3-th spectral band $\mathbf{F}^{(n_3)} = \mathcal{F}(:,:,n_3)$ is given by

$$\text{PSNR}_{n_3} = 10\log_{10}\left(\frac{\max(\mathbf{F}^{(n_3)})^2}{\text{MSE}(\mathbf{F}^{(n_3)}, \tilde{\mathbf{F}}^{(n_3)})}\right), \tag{22}$$

where $\max(\mathbf{F}^{(n_3)})$ is the maximum possible value of $\mathbf{F}^{(n_3)}$, and MSE is the mean squared error between $\mathbf{F}^{(n_3)}$, and $\tilde{\mathbf{F}}^{(n_3)}$.

B. Experiments

We evaluated the proposed algorithm in various scenarios. Tables I, II and III present the spatial PSNR when varying the noise levels, and the percentage of captured data for both databases, and both CSI architectures. Labels *DCT*, *KSVD*, and *JOINT* refer to the method used in the recovery process. DCT refers to the tt-SVD framework using only the DCT; KSVD refers to the tt-SVD using only the learned orthogonal transform described in [14], notice that, in this case, the transform domain is learned from the original data; and JOINT refers to the proposed method which jointly uses the DCT, and the learned transform from an initial estimation of the SI, respectively.

Table I shows the recovery results for the Jasper database acquired with CASSI when using only the DCT, and the orthogonal learned transform (KSVD). There, it is evident that insofar the noise level increases, the performance of the learned transform decays, while the DCT exhibits a stable behaviour.

Tables II and III show the comparison between the DCT and the JOINT method to recover the two databases $\mathcal{F}_1$, and $\mathcal{F}_2$, respectively . Notice that, for all cases, the joint method outperforms the results of using only the DCT even in the presence of noise.

TABLE I
RECOVERY QUALITY RESULTS IN TERMS OF SPATIAL PSNR [DB] USING CASSI VARYING THE NOISE LEVEL AND % OF CAPTURED DATA FOR $\mathcal{F}_2$ USING THE DCT AND THE KSVD APPROACHES

SRN [dB]	Approach	% of data			
		10	20	30	40
10	DCT	**18,130**	**17,623**	**16,584**	**15,664**
	KSVD	6,094	6,620	7,144	7,690
20	DCT	**22,376**	**23,915**	**24,210**	**24,215**
	KSVD	8,053	8,948	10,659	11,515
30	DCT	**23,568**	**26,233**	**28,170**	**29,234**
	KSVD	13,767	14,537	15,205	15,767
40	DCT	**23,686**	**26,770**	**28,991**	**31,219**
	KSVD	20,263	21,433	22,136	22,597
Inf	DCT	23,713	26,810	29,203	31,098
	KSVD	**25,923**	**29,376**	**31,927**	**34,413**

TABLE II
RECOVERY QUALITY RESULTS IN TERMS OF SPATIAL PSNR [DB] VARYING THE NOISE LEVEL AND % OF CAPTURED DATA FOR $\mathcal{F}_1$

System	**SNR [dB]**	Approach	% of data			
			10	20	30	40
CASSI	10	DCT	24,396	24,676	24,275	23,233
		JOINT	**27,368**	**26,688**	**25,234**	**24,677**
	20	DCT	24,921	25,925	27,173	27,752
		JOINT	**29,360**	**30,869**	**32,685**	**31,783**
	30	DCT	24,995	26,108	27,689	28,757
		JOINT	**30,175**	**32,617**	**35,080**	**36,421**
	40	DCT	25,004	26,132	27,757	28,895
		JOINT	**30,293**	**32,756**	**35,668**	**38,318**
	Inf	DCT	25,977	27,111	28,186	28,903
		JOINT	**30,363**	**32,983**	**36,072**	**38,478**
CCASSI	10	DCT	27,080	26,894	25,523	24,212
		JOINT	**29,751**	**27,945**	**25,669**	**24,216**
	20	DCT	27,929	31,380	32,545	32,527
		JOINT	**34,984**	**35,038**	**34,056**	**33,080**
	30	DCT	26,848	30,316	35,303	37,753
		JOINT	**37,794**	**40,579**	**42,073**	**42,018**
	40	DCT	26,853	29,240	33,571	36,263
		JOINT	**38,873**	**43,245**	**46,616**	**48,156**
	Inf	DCT	27,514	31,878	34,115	37,376
		JOINT	**40,979**	**47,694**	**53,212**	**55,905**

TABLE III
RECOVERY QUALITY RESULTS IN TERMS OF SPATIAL PSNR [DB] VARYING THE NOISE LEVEL AND % OF CAPTURED DATA FOR $\mathcal{F}_2$

System	**SNR [dB]**	Approach	% of data			
			10	20	30	40
CASSI	10	DCT	18,130	17,623	16,584	15,664
		JOINT	**20,897**	**21,839**	**21,887**	**21,062**
	20	DCT	22,376	23,915	24,210	24,215
		JOINT	**24,109**	**25,670**	**26,681**	**27,392**
	30	DCT	23,568	26,233	28,170	29,234
		JOINT	**25,307**	**28,605**	**29,750**	**31,360**
	40	DCT	23,686	26,770	28,991	31,219
		JOINT	**25,708**	**29,196**	**31,753**	**33,841**
	Inf	DCT	23,713	26,810	29,203	31,098
		JOINT	**25,404**	**29,544**	**31,747**	**33,927**
CCASSI	10	DCT	17,702	16,714	15,596	14,343
		JOINT	**20,081**	**21,265**	**20,052**	**19,943**
	20	DCT	23,274	24,214	24,057	23,423
		JOINT	**25,157**	**26,210**	**25,679**	**24,712**
	30	DCT	26,194	29,164	30,468	30,978
		JOINT	**30,266**	**32,815**	**33,488**	**33,242**
	40	DCT	26,837	31,054	33,788	35,781
		JOINT	**33,008**	**37,420**	**40,414**	**41,180**
	Inf	DCT	27,142	31,408	34,802	37,632
		JOINT	**33,331**	**41,592**	**47,060**	**47,947**

C. Analysis of the results

For most cases, the proposed method obtains the best performance. Particularly in the presence of noise it is evident that using only a learned transform is not sufficient to obtain an acceptable quality. In addition, the quality results of CCASSI were superior to those of CASSI. This behaviour is because CCASSI does not disperse nor multiplexes the spectral information, which allows the tensor to maintain a direct correspondence. Finally, the gain of including the Discrete Cosine transform jointly to a learned transform in the solution of the CSI recovery problem based on the tensor transform singular value decomposition gains up to 10dB in the presence of noise and up to 20 dB in the noiseless case.

On the other hand, the performance of the proposed method in presence of high levels of noise is better when there is lower amount of captured data. This is consistent with recovery problems because there is a relationship between the noise and the over-determined level of such problems, i.e., as more observations are acquired more over-determined is the problem which makes the solutions more sensible to noise.

VI. Conclusions and Future Work

This paper presented an three stage methodology to recover a SI from a set of compressed projections based on the tt-SVD. The main contribution is the formulation of an inverse problem which promotes a low-rank structure of the recovered SI in two domains: the DCT and a learned transform domain. This improves the recovery quality since the learned transform better represent the data, and provides robustness to the recovery process since the DCT is stable in the presence of noise. The transform was learned by using a fast sparsity-based orthogonal dictionary technique from an initial estimation of the scene. The inverse problem was solved by following the extended ADMM approach. We tested the performance of the proposal varying the optical system, the amount of captured data, and the noise level. Simulations in two databases show that the method improves the quality of the reconstructions in terms of the spatial PSNR in up to 10dB in the presence of noise, and in up to 20dB in the noiseless cases.

References

[1] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 105–115, 2014.

[2] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 83–91, 2008.

[3] J. Bacca, H. Vargas, and H. Arguello, "A constrained formulation for compressive spectral image reconstruction using linear mixture models," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2017 IEEE 7th International Workshop on*, pp. 1–5, IEEE, 2017.

[4] N. Renard and S. Bourennane, "Dimensionality reduction based on tensor modeling for classification methods," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 4, p. 1123, 2009.

[5] J. M. Bioucas-Dias and J. M. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, 2008.

[6] G.-J. Song, X. Zhang, Q. Jiang, and M. K. Ng, "Robust tensor completion using transformed tensor singular value decomposition,"

[7] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 1, pp. 148–172, 2013.

[8] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 208–220, 2013.

[9] R. A. Harshman, "Foundations of the parafac procedure: Models and conditions for an" explanatory" multimodal factor analysis," *UCLA Working Papers in Phonetics*, vol. 1, pp. 1 – 84, 1970.

[10] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.

[11] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra and its Applications*, vol. 435, no. 3, pp. 641–658, 2011.

[12] H. Palangi, A. Ghafari, M. Babaie-Zadeh, and C. Jutten, "Image coding and compression with sparse 3d discrete cosine transform," in *International Conference on Independent Component Analysis and Signal Separation*, pp. 532–539, Springer, 2009.

[13] C. V. Correa-Pugliese, D. F. Galvis-Carreño, and H. Arguello-Fuentes, "Sparse representations of dynamic scenes for compressive spectral video sensing," *Dyna*, vol. 83, no. 195, pp. 42–51, 2016.

[14] C. Bao, J.-F. Cai, and H. Ji, "Fast sparsity-based orthogonal dictionary learning for image restoration," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3384–3391, 2013.

[15] C. V. Correa, H. Arguello, and G. R. Arce, "Compressive spectral imaging with colored-patterned detectors," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pp. 7789–7793, IEEE, 2014.

[16] X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, L. Carin, and D. J. Brady, "Computational snapshot multispectral cameras: Toward dynamic capture of the spectral world," *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 95–108, 2016.

[17] E. Kernfeld, M. Kilmer, and S. Aeron, "Tensor–tensor products with invertible linear transforms," *Linear Algebra and its Applications*, vol. 485, pp. 545–570, 2015.

[18] J. Yang and X. Yuan, "Linearized augmented lagrangian and alternating direction methods for nuclear norm minimization," *Mathematics of computation*, vol. 82, no. 281, pp. 301–329, 2013.

[19] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[20] C. Chen, B. He, Y. Ye, and X. Yuan, "The direct extension of admm for multi-block convex minimization problems is not necessarily convergent," *Mathematical Programming*, vol. 155, no. 1-2, pp. 57–79, 2016.

[21] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, p. 2010, 1956.

[22] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and bregman iterative methods for matrix rank minimization," *Mathematical Programming*, vol. 128, no. 1-2, pp. 321–353, 2011.

[23] F. Zhu, Y. Wang, B. Fan, S. Xiang, G. Meng, and C. Pan, "Spectral unmixing via data-guided sparsity," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5412–5427, 2014.