HYPERSPECTRAL IMAGE INPAINTING BASED ON COLLABORATIVE TOTAL VARIATION

P. Addesso[1], M. Dalla Mura[2], L. Condat[2], R. Restaino[1], G. Vivone[1], D. Picone[1], J. Chanussot[2,3]

[1]DIEM, University of Salerno, Italy.
[2]GIPSA-Lab, CNRS, University of Grenoble Alpes, Grenoble, France.
[3]Faculty of Electrical and Computer Engineering, University of Iceland, Iceland.

ABSTRACT

Inpainting in hyperspectral imagery is a challenging research area and several methods have been recently developed to deal with this kind of data. In this paper we address missing data restoration via a convex optimization technique with regularization term based on Collaborative Total Variation (CTV). In particular we evaluate the effectiveness of several instances of CTV in conjunction with different dimensionality reduction algorithms.

Index Terms— Inpainting, Hyperspectral images, Total Variation, Collaborative Norms, Image restoration.

1. INTRODUCTION

HyperSpectral (HS) images are composed of hundreds of spectral channels, typically ranging from the visible to the long-wave infrared. The fine sampling of the spectral domain provided by HS imaging systems allows one to infer the spectral signature of the materials present in a scene, as required by several applications such as Earth observation (e.g., precision agriculture, land cover classification and mineralogy mapping [1]), detection of chemical agent plumes [2], quality control [3], astrophysics [4], biometrics [5], archeology [6] and biomedical imaging [7].

Hyperspectral acquisition can be affected by missing information due to sensors defects (e.g., damaged pixels, striping and sensor saturation), presence of occlusions (such as clouds in remote sensed images [8]) and missing data (e.g., due to image reprojection or incomplete spatial coverage of the acquisition). In some other cases, the imaging system is specifically designed to acquire images with missing information such as in compressed measurements schemes [9, 10]. In order to deal with the problem of missing data reconstruction, several techniques have been proposed ranging from the *diffusion-based* methods to variational approaches [11, 12]. These latter are typically casted as *convex optimization* problems adopting spatial regularization, e.g. via Total Variation (TV) [13, 14], which enforces piece-wise smoothness in the reconstructed images. The definition of TV for vectorial images (e.g., RGB) is not unique and affects how image discontinuities (i.e., edges) are coupled across bands [15]. This aspect is even more relevant when dealing with HS data due to the high number of spectral channels. This work is specifically devoted to explore several vectorial formulations of TV based on the *Collaborative Total Variation (CTV)* paradigm [16] for the inpainting of HS images. In particular, we will analyze several CTV-based regularization strategies and different architectures to cope with the high dimensionality of the HS data.

2. COLLABORATIVE TOTAL VARIATION

We denote a HS image $\mathbf{H}$ in vector notation: $\mathbf{H} \in \mathbb{R}^{C \times N}$, with C bands and N pixels (each row contains the lexicographically ordered pixels of a given band). In this scenario, it is possible to define the so-called *inpainting domain* $\mathcal{I} \subset \mathbb{R}^N$, i.e. the set of all the pixels in the image in which the spectral information is completely missing. Performing the inpainting is an ill-posed inverse problem that can be addressed by a convex optimization formulation:

$$\underset{\mathbf{Z}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{H} - \mathbf{Z}\|^2_{F(\overline{\mathcal{I}})} + \lambda_\varphi \varphi(\mathbf{Z}), \qquad (1)$$

where *i)* $\mathbf{Z} \in \mathbb{R}^{C \times N}$ is the image to estimate, *ii)* $\|\cdot\|_{F(\overline{\mathcal{I}})}$ is the Frobenius norm computed on the complement to the inpainting domain $\mathcal{I}$; *iii)* $\varphi(\mathbf{Z})$ is a regularization term accounting for the desired features of the solution, whose weight (or regularization parameter) λ_φ has to be tuned.

Since $\mathbf{Z}$ lives in a subspace of dimensionality (significantly) lower than C [17], it is possible to solve an alternative problem which applies on the data after dimensionality reduction. For example, $\mathbf{Z}$ can be factorized as $\mathbf{Z} = \mathbf{E}\mathbf{X}$, in which $\mathbf{E}$ is the set of basis (with cardinality $L \leq C$) spanning the subspace of $\mathbf{Z}$ and $\mathbf{X}$ are the representation coefficients. Therefore the problem becomes

$$\underset{\mathbf{X}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{H} - \mathbf{E}\mathbf{X}\|^2_{F(\overline{\mathcal{I}})} + \lambda_\varphi \varphi(\mathbf{X}). \qquad (2)$$

In the following we will focus on the regularization term, in order to identify the most appropriate CTV instance for HS

Table 1: Pavia University: numerical results computed via an Intel Core i7 vPro laptop. The best overall results for each inpainting domain pattern are in blue-bold. The best results within every case are in bold and the second-best ones are underlined. The indices have been computed by averaging over 10 Monte Carlo trials for SVD ($L = 10$) and NF and 50 Monte Carlo trials for VCA ($L = 10$), with $SNR = 30$ dB. PSNR and SAM standard deviation is $< 5 \cdot 10^{-4}$ for SVD and NF, and it is $< 2 \cdot 10^{-2}$ for VCA. UIQI standard deviation is $< 2 \cdot 10^{-5}$ for SVD and NF, and it is $< 2 \cdot 10^{-4}$ for VCA.

	Algorithm	λ_φ	Random			Stripes ($s=1$)			Stripes ($s=2$)			T (s)
			PSNR	UIQI	SAM	PSNR	UIQI	SAM	PSNR	UIQI	SAM	
VCA	CTV-$\ell^{2,2,1}(dbx)$	1e-2	**30.93**	**0.955**	2.661	**31.87**	**0.963**	2.611	29.20	0.932	**3.153**	22.92
	CTV-$\ell^{1,1,1}(bdx)$	1e-2	29.76	0.942	2.951	29.94	0.942	2.958	27.30	0.899	3.568	23.27
	CTV-$\ell^{2,1,1}(bdx)$	1e-3	30.33	0.948	2.617	31.32	0.957	**2.528**	28.00	0.912	3.335	23.37
	CTV-$\ell^{\infty,1,1}(bdx)$	5e-2	30.13	0.945	3.064	30.70	0.951	3.086	27.84	0.908	3.727	27.08
	CTV-$\ell^{\infty,\infty,1}(bdx)$	1e-1	29.76	0.942	3.264	30.70	0.953	3.196	28.51	0.925	3.658	27.97
	CTV-$\ell^{2,\infty,1}(dbx)$	1e-1	30.29	0.947	3.207	30.81	0.953	3.221	28.59	0.923	3.706	24.88
	CTV-$(\mathbb{S}^1(bd),\ell^1(x))$	2e-2	30.34	0.947	2.860	31.05	0.954	2.772	28.91	0.925	3.268	25.74
	CTV-$(\mathbb{S}^\infty(bd),\ell^1(x))$	1e-2	30.63	0.953	2.757	31.60	0.962	2.730	**29.32**	**0.936**	3.231	85.60
SVD	CTV-$\ell^{2,2,1}(dbx)$	5e-3	**30.97**	**0.955**	**2.526**	31.80	**0.962**	2.494	**29.22**	**0.932**	3.025	20.38
	CTV-$\ell^{1,1,1}(bdx)$	2e-2	29.96	0.942	3.229	30.65	0.949	3.231	27.89	0.907	3.862	19.92
	CTV-$\ell^{2,1,1}(bdx)$	5e-3	30.14	0.948	2.612	31.16	0.957	2.575	28.16	0.916	3.285	20.35
	CTV-$\ell^{\infty,1,1}(bdx)$	2e-2	29.88	0.945	2.885	30.74	0.952	2.867	27.78	0.908	3.580	22.86
	CTV-$\ell^{\infty,\infty,1}(bdx)$	2e-2	29.85	0.944	2.960	30.81	0.954	2.867	28.47	0.924	3.417	24.45
	CTV-$\ell^{2,\infty,1}(dbx)$	1e-2	30.44	0.950	2.850	31.10	0.956	2.839	28.66	0.926	3.403	22.27
	CTV-$(\mathbb{S}^1(bd),\ell^1(x))$	5e-3	30.69	0.951	2.580	31.49	0.958	**2.486**	29.13	0.929	**2.990**	22.85
	CTV-$(\mathbb{S}^\infty(bd),\ell^1(x))$	5e-3	30.72	0.953	2.744	31.48	0.960	2.733	29.09	0.931	3.175	81.42
NF	CTV-$\ell^{2,2,1}(dbx)$	5e-3	**30.86**	**0.953**	**3.227**	**31.64**	**0.960**	3.197	**29.12**	**0.930**	3.693	385.67
	CTV-$\ell^{1,1,1}(bdx)$	5e-3	29.88	0.941	3.444	30.53	0.946	3.420	27.91	0.906	4.029	374.82
	CTV-$\ell^{2,1,1}(bdx)$	5e-3	30.13	0.947	3.255	31.11	0.955	3.200	28.11	0.914	3.887	394.85
	CTV-$\ell^{\infty,1,1}(bdx)$	2e-1	29.82	0.940	3.370	30.54	0.948	3.371	27.74	0.905	4.006	461.96
	CTV-$\ell^{\infty,\infty,1}(bdx)$	2e-1	29.61	0.939	3.556	30.61	0.951	3.491	28.43	0.920	3.966	467.85
	CTV-$\ell^{2,\infty,1}(dbx)$	2e-1	30.08	0.943	3.514	30.65	0.950	3.526	28.42	0.917	3.995	427.67
	CTV-$(\mathbb{S}^1(bd),\ell^1(x))$	5e-3	30.57	0.949	3.254	31.35	0.957	**3.145**	29.00	0.927	**3.620**	429.26
	CTV-$(\mathbb{S}^\infty(bd),\ell^1(x))$	5e-3	30.56	0.950	3.561	31.23	0.957	3.569	28.98	0.929	3.936	636.34

inpainting. More in detail, we exploit the so-called *Collaborative Norms* (CNs), which are denoted by $\|\cdot\|_C$ and applied to a multivariate image $\mathbf{A} \in \mathbb{R}^{L \times N \times M}$, where L is the number of the basis vectors, N is the number of the pixels and M is the number of directional derivatives computed on each pixel.

There are two general formulations for the CN's. The first one is based on the $\ell^{p,q,r}$ norm. For example, if we associate the ℓ^p norm to the derivative dimension (d), the ℓ^q norm to the basis dimension (b) and the ℓ^r norm to the pixel dimension (x), i.e. the *application order* is (dbx), we obtain the $\ell^{p,q,r}(dbx)$ norm:

$$||\mathbf{A}||_{p,q,r} = \left[\sum_{j=1}^{N} \left[\sum_{i=1}^{L} \left[\sum_{k=1}^{M} |A_{i,j,k}|^p \right]^{q/p} \right]^{r/q} \right]^{1/r}. \quad (3)$$

The other formulation is based on the Schatten p-norm (referred to as $\mathbb{S}^p$) applied to two dimensions (usually spectra and derivatives) and then on the ℓ^q norm for the third dimension (i.e. the pixel dimension), according to:

$$(\mathbb{S}^p, \ell^q)(\mathbf{A}) = \left[\sum_{j=1}^{N} \left[\sum_{t=1}^{\min\{L,M\}} \sigma_t^p(\mathbf{A}_{\cdot,j,\cdot}) \right]^{q/p} \right]^{1/q}, \quad (4)$$

that defines the $(\mathbb{S}^p(bd), \ell^q(x))$ norm, where $\mathbf{A}_{\cdot,j,\cdot} \in \mathbb{R}^{L \times M}$ is the j-th two-dimensional matrix and $\sigma_t(\cdot)$ are its *singular values* (e.g., see [18]). In our framework $\mathbf{A}$ can be written as $\mathbf{A}(\mathbf{X}) = \text{cat}_3[\mathbf{X}\mathbf{D}_h, \mathbf{X}\mathbf{D}_v]$, where the operator $\text{cat}_d[\cdot, \cdot]$ concatenates two matrices along the direction d and where $\mathbf{X}\mathbf{D}_h$ and $\mathbf{X}\mathbf{D}_v$ stand for the spatial derivatives of $\mathbf{X}$ in the horizontal and vertical directions, respectively. Thus, the regularization term can be recast as $\varphi(\mathbf{X}) = \|\mathbf{A}(\mathbf{X})\|_C$. The range of possible norms is wide [16], due to the possibility of freely selecting the norm kind and the application order. Therefore, we test the effectiveness of the most popular choices, such as the $\ell^{2,1}(dbx)$ norm and the $(\mathbb{S}^1(bd), \ell^1(x))$, with or without the application of a dimensionality reduction technique.

3. ALGORITHM DETAILS

In order to solve the problem in eq.(2), we use the alternating direction method of multipliers (ADMM) approach [19], that relies on the introduction of auxiliary variables into the optimization problem, via the so-called variable splitting method. More in detail, we split the original optimization variable $\mathbf{X}$ into a total of four variables, i.e. $\mathbf{X}$ itself and three auxiliary variables, i.e.,$\mathbf{U}_1$, $\mathbf{U}_2$ and $\mathbf{U}_3$. Therefore the optimization
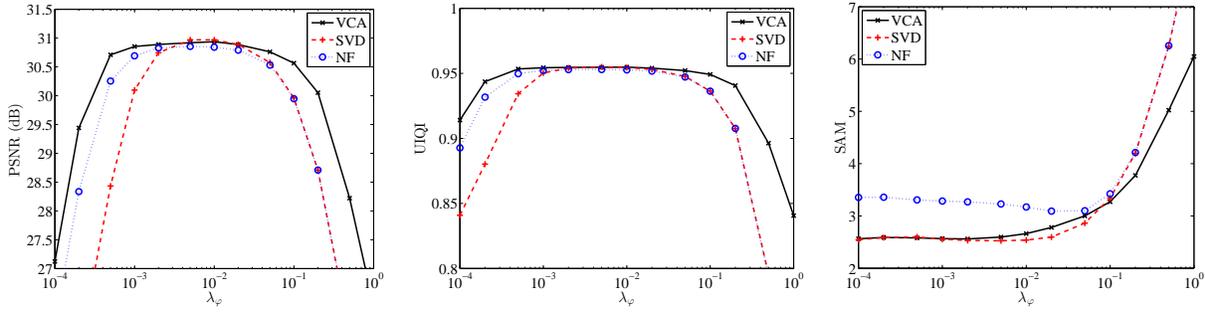
Fig. 1: Quality indexes computed with respect to λ_φ using CTV-$\ell^{2,2,1}(dbx)$ algorithm (Random pattern). Left plot: PSNR. Central plot: UIQI. Right plot: SAM. A factorization with $L = 10$ has been performed when using VCA and SVD.

problem can be written as:

$$\underset{\mathbf{X},\mathbf{U}_1,\mathbf{U}_2,\mathbf{U}_3}{\text{minimize}} \quad \tfrac{1}{2}\left\|\mathbf{H} - \mathbf{E}\mathbf{U}_1\right\|_{F(\overline{\mathcal{I}})}^2 + \lambda_\varphi \varphi(\mathbf{U}_2, \mathbf{U}_3),$$

$$\text{subject to: } \mathbf{U}_1 = \mathbf{X}; \ \mathbf{U}_2 = \mathbf{X}\mathbf{D}_h; \ \mathbf{U}_3 = \mathbf{X}\mathbf{D}_v.$$
(5)

By introducing also the scaled dual variables $\mathbf{V}_1, \mathbf{V}_2$ and $\mathbf{V}_3$, we can use an augmented Lagrangian formulation and it is possible to write the algorithm described below.

Initialization: $\mathbf{X}^{(0)} = \mathbf{0}, \mathbf{U}_a^{(0)} = \mathbf{0}, \mathbf{V}_a^{(0)} = \mathbf{0}, \forall a \in \{1,2,3\}$, where $\mathbf{0}$ is a null matrix.

Repeat the following steps:

Step 1: compute $\mathbf{X}^{(k+1)}$.

$$\begin{aligned}
\mathbf{X}^{(k+1)} &= \left[\mathbf{I} + \mathbf{D}_h\mathbf{D}_h^T + \mathbf{D}_v\mathbf{D}_v^T\right]^{-1} \\
&\times \Big[\left(\mathbf{U}_1^{(k)} + \mathbf{V}_1^{(k)}\right) + \left(\mathbf{U}_2^{(k)} + \mathbf{V}_2^{(k)}\right)\mathbf{D}_h^T \\
&\quad + \left(\mathbf{U}_3^{(k)} + \mathbf{V}_3^{(k)}\right)\mathbf{D}_v^T\Big],
\end{aligned}$$

where $\mathbf{I}$ is the identity matrix.

Step 2: compute $\mathbf{U}_1^{(k+1)}$.

$$\begin{aligned}
\mathbf{U}_1^{(k+1)} \circ \mathbf{M} &= \left[\mathbf{E}\mathbf{E}^T - \mu\mathbf{I}\right]^{-1} \\
&\times \left[\mathbf{E}^T\mathbf{H} + \mu\left(\mathbf{X}^{(k+1)} - \mathbf{V}_1^{(k)}\right)\right] \circ \mathbf{M} \\
\mathbf{U}_1^{(k+1)} \circ \overline{\mathbf{M}} &= \left(\mathbf{X}^{(k+1)} - \mathbf{V}_1^{(k)}\right) \circ \overline{\mathbf{M}},
\end{aligned}$$

where μ is the *penalty parameter* introduced in the augmented Lagrangian formulation (in our setup $\mu = 0.05$), $\mathbf{M}$ is a mask whose elements are zero on the inpainting domain and one otherwise, and the operator "$\circ$" indicates the element-wise Hadamard product.

Step 3: compute $\mathbf{U}_2^{(k+1)}$ and $\mathbf{U}_3^{(k+1)}$.

$$\left\{\mathbf{U}_2^{(k+1)}, \mathbf{U}_3^{(k+1)}\right\} = \text{prox}_\varphi\left(\mathbf{C}, \frac{\lambda_\varphi}{\mu}\right),$$

where

$$\mathbf{C} = \left\{\left(\mathbf{X}^{(k+1)}\mathbf{D}_h - \mathbf{V}_2^{(k)}\right), \left(\mathbf{X}^{(k+1)}\mathbf{D}_v - \mathbf{V}_3^{(k)}\right)\right\}$$

and $\text{prox}_\varphi(\cdot,\cdot)$ is the *proximity operator* related to the norm used for the regularization term $\varphi(\cdot)$.

Step 4: update the Lagrange multipliers.

$$\begin{aligned}
\mathbf{V}_1^{(k+1)} &= \mathbf{V}_1^{(k)} - \left(\mathbf{X}^{(k+1)} - \mathbf{U}_1^{(k+1)}\right) \\
\mathbf{V}_2^{(k+1)} &= \mathbf{V}_2^{(k)} - \left(\mathbf{X}^{(k+1)}\mathbf{D}_h - \mathbf{U}_2^{(k+1)}\right) \\
\mathbf{V}_3^{(k+1)} &= \mathbf{V}_3^{(k)} - \left(\mathbf{X}^{(k+1)}\mathbf{D}_v - \mathbf{U}_3^{(k+1)}\right)
\end{aligned}$$

Until a suitable stopping criterion is satisfied.

4. NUMERICAL RESULTS

The effect of selecting different CNs is evaluated on a real HS image acquired by the ROSIS sensor over the city of Pavia, Italy (103 spectral channels at 1.3 m of spatial resolution). The inpainting domain $\mathcal{I}$ has been constructed following two strategies: *a*) *Random pattern*; *b*) *Stripes pattern* of width s, which is a typical defect affecting remotely sensed images acquired by some sensors [8]. In this paper we considered missing data corresponding to 50% of the pixels, either organized in a random or stripes ($s = \{1, 2\}$) patterns. White Gaussian noise was added to the images (to mimic different acquisitions on the same area) leading to a $SNR = 30$ dB.

The numerical results will be devoted to test the effectiveness of the Collaborative Norms for image inpainting in these scenarios: *i*) factorization via Vertex Component Analysis (VCA); *ii*) factorization via Singular Value Decomposition (SVD); *iii*) No Factorization (NF), i.e. $L = C$ and $\mathbf{E} = \mathbf{I}$.

The inpainting results are evaluated via several indices: the Peak Signal-to-Noise Ratio (PSNR), the Universal Image Quality Index (UIQI) [20] and the Spectral Angle Mapper (SAM) [21] (useful to evaluate the spectral distortion, which is of paramount importance for HS data). By a preliminary Monte Carlo simulation phase, we found that reasonable results are typically obtained when $L \in [10, 20]$. Therefore we select the value $L = 10$ for our experiments, for both VCA and SVD factorization, leading to a computational time less than half with respect to $L = 20$.
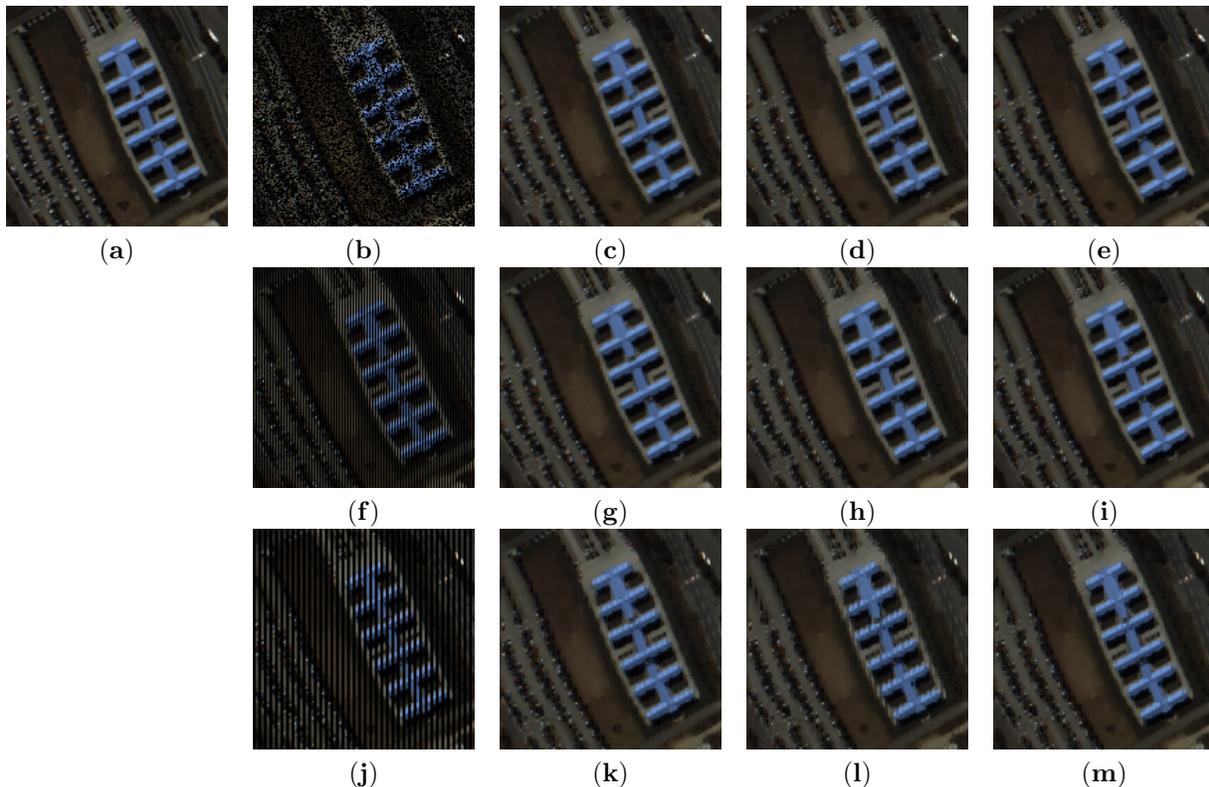
Fig. 2: Pavia University. Panel (a): Original image. *Panels* (b) - (e): *Random pattern*. Degraded images (b), estimated image via VCA/CTV: $\ell^{2,2,1}(dbx)$ (c), estimated image via VCA/CTV: $\ell^{1,1,1}(bdx)$ (d), estimated image via VCA/CTV: $(\mathbb{S}^1(bd),\ell^1(x))$ (e). *Panels* (f) - (i): *Stripes pattern* ($s=1$). Degraded images (f), estimated image via VCA/CTV: $\ell^{2,2,1}(dbx)$ (g), estimated image via VCA/CTV: $\ell^{1,1,1}(bdx)$ (h), estimated image via VCA/CTV: $(\mathbb{S}^1(bd),\ell^1(x))$ (i). *Panels* (j) - (m): *Stripes pattern* ($s=2$). Degraded images (j), estimated image via VCA/CTV: $\ell^{2,2,1}(dbx)$ (k), estimated image via VCA/CTV: $\ell^{1,1,1}(bdx)$ (l), estimated image via VCA/CTV: $(\mathbb{S}^1(bd),\ell^1(x))$ (m). Other relevant parameters are $L=10$ and $SNR=30$ dB.

The results are reported in Table 1 (the regularization coefficient λ_φ was selected by optimizing the PSNR values). It is evident that the best norms for CTV, regardless of the used factorization, are $\ell^{2,2,1}(dbx)$ (overall the best one), $(\mathbb{S}^\infty(bd),\ell^1(x))$ and $(\mathbb{S}^1(bd),\ell^1(x))$. This can be explained because the ℓ^2 and the Schatten norms applied on the bands lead to a coupling of spatial discontinuities across channels, e.g., by enforcing smooth transitions among adjacent bands as for ℓ^2. Regularizing with norms such as $\ell^{1,1,1}(bdx)$, $\ell^{\infty,1,1}(bdx)$ and $\ell^{\infty,\infty,1}(bdx)$ have led to poor performance, which can be due to a "too weak" (ℓ^1) or "too strong" (ℓ^∞) spectral coupling of edges.

If we focus on the dimensionality reduction, the results show that it is effective for inpainting, both in terms of accuracy in the reconstructed image (SAM dramatically improves) and in reducing the computational burden. In particular SVD seems to have a little advantage over VCA in spectral fidelity. On the other hand, as shown in Fig. 1, VCA exhibits a higher degree of robustness with respect to the choice of λ_φ, at least for PSNR and UIQI. This trend, which is shown here only in the case of $\ell^{2,2,1}(dbx)$ norm, holds also for the other CNs.

Finally, we notice that the algorithms perform well when the stripes are thin ($s=1$), but the performance strongly decreases when they are thicker ($s=2$) due to limitations of the local approach. This effect can be appreciated also by a visual analysis. Indeed in Fig. 2 it is clear to see the artifacts due to thick stripes, especially for anisotropic norms such as $\ell^{1,1,1}(dbx)$.

5. CONCLUSION

In this work we compared the effectiveness of several norms to implement the Collaborative Total Variation framework for the inpainting of HS images affected by damaged pixels and stripes. The numerical results show that that performing a dimensionality reduction and using norms $\ell^{2,2,1}(dbx)$, $\mathbb{S}^\infty(bd),\ell^1(x)$ and $\mathbb{S}^1(bd),\ell^1(x)$ allows to obtain good performance, above all in reconstructing the spectral information. Future work will be devoted to the extension of this framework to the case of Non-Local TV, that has been proved to be very effective (e.g. see [22]), and to the case of multitemporal reconstruction [8].

6. REFERENCES

[1] J. M. Bioucas-Dias, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 6–36, 2013.

[2] D. Manolakis, S. Golowich, and R. S. DiPietro, "Longwave infrared hyperspectral remote sensing of chemical clouds: A focus on signal processing approaches," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 120–141, July 2014.

[3] AA Gowen, CPo O'Donnell, PJ Cullen, G Downey, and JM Frias, "Hyperspectral imaging–an emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology*, vol. 18, no. 12, pp. 590–598, 2007.

[4] E. Villeneuve and H. Carfantan, "Nonlinear deconvolution of hyperspectral data with mcmc for studying the kinematics of galaxies," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4322–4335, Oct 2014.

[5] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition with spatiospectral information fusion and pls regression," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 1127–1137, March 2015.

[6] Haida Liang, "Advances in multispectral and hyperspectral imaging for archaeology and art conservation," *Applied Physics A*, vol. 106, no. 2, pp. 309–323, 2012.

[7] Tuan Vo-Dinh, "A hyperspectral imaging system for in vivo optical diagnostics," *IEEE Engineering in Medicine and Biology Magazine*, vol. 23, no. 5, pp. 40–49, 2004.

[8] H. Shen, X. Li, Q. Cheng, C. Zeng, G. Yang, H. Li, and L. Zhang, "Missing information reconstruction of remote sensing data: A technical review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 3, no. 3, pp. 61–85, Sept 2015.

[9] C. Li, T. Sun, K. F. Kelly, and Y. Zhang, "A compressive sensing and unmixing scheme for hyperspectral data processing," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1200–1210, March 2012.

[10] D. T. Eason and M. Andrews, "Total variation regularization via continuation to recover compressed hyperspectral images," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 284–293, Jan 2015.

[11] C. Guillemot and O. Le Meur, "Image inpainting : Overview and recent advances," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 127–144, Jan 2014.

[12] J. Zhang, D. Zhao, R. Xiong, S. Ma, and W. Gao, "Image restoration using joint statistical modeling in a space-transform domain," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 6, pp. 915–928, June 2014.

[13] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1852–1866, June 2015.

[14] W. He, H. Zhang, L. Zhang, and H. Shen, "Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 1, pp. 178–188, Jan 2016.

[15] P. Blomgren and T. F. Chan, "Color tv: total variation methods for restoration of vector-valued images," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 304–309, Mar 1998.

[16] J. Duran, M. Moeller, C. Sbert, and D. Cremers, "Collaborative total variation: A general framework for vectorial tv models," *SIAM Journal on Imaging Sciences*, vol. 9, no. 1, pp. 116–151, 2016.

[17] K. Cawse-Nicholson, S. Damelin, A. Robin, and M. Sear, "Determining the intrinsic dimension of a hyperspectral image using random matrix theory," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1301–1310, Apr. 2013.

[18] E. Deza and M.M. Deza, *Dictionary of Distances*, Elsevier, Amsterdam, NL, 1st edition, 2006.

[19] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "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, Jan 2011.

[20] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[21] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci.Workshop*, 1992, pp. 147–149.

[22] G. Chierchia, N. Pustelnik, B. Pesquet-Popescu, and J. C. Pesquet, "A nonlocal structure tensor-based approach for multicomponent image recovery problems," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5531–5544, Dec 2014.