Reconstruction From Random Projections of Hyperspectral Imagery With Spectral and Spatial Partitioning — Source link ↗

Nam Hoai Ly, Qian Du, James E. Fowler

Institutions: Mississippi State University

Related papers:

- Compressed sensing

- Integration of Spectral–Spatial Information for Hyperspectral Image Reconstruction From Compressive Random Projections

- An Introduction To Compressive Sampling

- Compressive-Projection Principal Component Analysis

- Learning recovered pattern from region-dependent model for hyperspectral imagery

Reconstruction From Random Projections of Hyperspectral Imagery With Spectral and Spatial Partitioning

Nam Hoai Ly, *Student Member, IEEE*, Qian Du, *Senior Member, IEEE*, and James E. Fowler, *Senior Member, IEEE*

Abstract—Random projections have recently been proposed to enable dimensionality reduction in resource-constrained sensor devices such that the computational burden is shifted to the receiver side of the system in the form of a reconstruction process. While a number compressed-sensing algorithms can provide such reconstruction, the principal-component based compressive-projection principal component analysis (CPPCA) algorithm has been shown to offer better performance for hyperspectral imagery. CPPCA is extended to incorporate both spectral and spatial partitioning of the hyperspectral dataset with experimental results evaluating reconstruction quality not only in terms of squared-error and spectral-angle fidelity but also via performance of the reconstructed data in classification and unmixing tasks. While experimental results demonstrate that either form of partitioning yields significantly better reconstruction than the original, non-partitioned algorithm, CPPCA using both spectral and spatial partitioning together outperforms either of the two used alone.

Index Terms—Hyperspectral imagery, principal component analysis, random projection.

I. INTRODUCTION

THERE has been significant interest in recent literature focused on dimensionality reduction using random projections [1]–[3]. It is anticipated that this random-projection process would not be explicitly calculated but, instead, implemented directly within the hardware of a sensing device, thereby providing dimensionality reduction simultaneously with signal acquisition. Thus, computational burden would shift from a resource-constrained sensor device to a reconstruction process implemented on a more powerful receiving device. Significant attention in recent literature has been devoted to the design of algorithms to reconstruct datasets from random projections.

In this paper, we are concerned specifically with the reconstruction from random projections of hyperspectral imagery. Consider an N-band hyperspectral dataset of consisting M vectors $\mathbf{X} = [\ \mathbf{x}_1 \ \cdots \ \mathbf{x}_M\ ]$, wherein each spectral pixel vector $\mathbf{x}_m \in \mathbb{R}^N$. The sensor produces random projections of these

vectors in the form of $\tilde{\mathbf{y}}_m = \mathbf{P}^T \mathbf{x}_m$, where randomly-chosen $N \times K$ matrix $\mathbf{P}$ forms an orthogonal projection onto a K-dimensional subspace. The problem at hand is then to design a reconstruction process at the receiver that recovers the dataset $\mathbf{X}$ from the projected vectors $\tilde{\mathbf{Y}} = [\ \tilde{\mathbf{y}}_1 \ \cdots \ \tilde{\mathbf{y}}_M\ ]$. The problem is often recast as recovering transform coefficients $\check{\mathbf{X}}$ for a transform basis $\mathbf{\Psi}$ from random projections $\tilde{\mathbf{Y}}$ such that

$$\tilde{\mathbf{Y}} = \mathbf{P}^T \mathbf{X} = \mathbf{P}^T \mathbf{\Psi} \check{\mathbf{X}}. \tag{1}$$

Typically, the performance of reconstruction algorithms depends on the degree of dimensionality reduction inherent in the random projections; this quantity is characterized as K/N, the *subrate*, of the random projections.

There are several approaches that one can take to perform a reconstruction of $\mathbf{X}$ from $\tilde{\mathbf{Y}}$. The most widely known is perhaps compressed sensing (CS); see [4] for an overview. However, [1] proposes an alternative in which reconstruction is based on principal component analysis (PCA) performed within the receiving device. Effectively, this compressive-projection PCA (CPPCA) recovers, solely from the random projections, approximations to the PCA eigenvectors of the dataset, using this recovered PCA basis to reconstruct the dataset. Experimental results in [1] demonstrated that, for hyperspectral data, CPPCA yielded much higher-quality reconstructions than CS in the terms of signal-to-noise ratio (SNR) of the reconstruction at a fixed subrate.

While CPPCA was originally applied in [1] to an entire hyperspectral dataset, recent work has shown that reconstruction quality can be further improved by having the receiver first segment the random projections into multiple partitions consisting of highly correlated data with the partitions being reconstructed independently. For example, [2] segments a hyperspectral image into normal and anomaly pixel classes, reconstructing each pixel class separately. Similarly, [5] employs several forms of supervised and unsupervised classification to spatially partition the dataset into multiple classes, applying CPPCA reconstruction to each spatial class independently.

In this paper, we extend this concept of partitioned reconstruction to the spectral dimension of a hyperspectral dataset. In essence, we incorporate into CPPCA the paradigm of segmented PCA (SPCA) [6], [7] in which a dataset is segmented into multiple spectral partitions, and PCA is applied in each segment independently. Because its eigendecompositions are limited to spectrally neighboring bands which are highly correlated, SPCA tends to outperform traditional PCA. The primary contribution of this paper is thus the extension of these SPCA advantages to

CPPCA. Additionally, we couple the classification-driven spatial partitioning proposed in [5] to this spectral partitioning, producing a spectral-spatial partitioned CPPCA variant. We find that the combination of spectral and spatial partitioning of this latter method outperforms not only CPPCA using either form of partitioning alone, but several forms of CS-based reconstruction as well.

The remainder of the discussion is organized as follows. We first briefly overview the process of dimensionality reduction via random projection as well as CPPCA and CS reconstruction from such random projections in Section II. Next, we present the spectrally partitioned and spectral-spatial partitioned CPPCA approaches in Sections III and IV, following with experimental results in Section V. Finally, several concluding remarks are made in Section VI.

II. RECONSTRUCTION FROM RANDOM PROJECTIONS

A. CS

There are numerous CS-based algorithms that can reconstruct $\mathbf{X}$ from $\tilde{\mathbf{Y}}$. Most of these CS algorithms recover a single vector at a time; for a dataset of multiple, possibly correlated vectors such as arising when $\mathbf{X}$ is a hyperspectral image, a simultaneous, or "multitask," reconstruction—such as multitask Bayesian CS (MT-BCS) [8]—is more appropriate.

Inherent to CS reconstruction is a basis which is assumed to yield a sparse representation of the signal to be reconstructed. Most frequently, a fixed, known basis $\boldsymbol{\Psi}$, such as a discrete wavelet transform (DWT), is used. For example, MT-BCS reconstruction is then

$$\check{\mathbf{X}} = \text{MT-BCS}(\mathbf{P}, \tilde{\mathbf{Y}}, \boldsymbol{\Psi}). \tag{2}$$

Alternatively, if the reconstruction process has access to a suitable body of (unprojected) training data, then a variety of dictionary-learning strategies can be applied to train the sparsity basis; often, such a learned basis will be overcomplete. For CS reconstruction, the learned dictionary $\boldsymbol{\Psi}$ is used as the sparsity basis for some suitable CS algorithm. For example, [9] couples learning of an overcomplete dictionary using the K-SVD learning algorithm [9] with orthogonal matching pursuit (OMP) [10]—we call the resulting CS reconstruction K-SVD+OMP,

$$\check{\mathbf{X}} = \text{OMP}(\mathbf{P}, \tilde{\mathbf{Y}}, \boldsymbol{\Psi}). \tag{3}$$

In subsequent experimental results, we use both the CS reconstructions of (2) and (3).

B. CPPCA

CPPCA [1] has been shown not only to outperform CS reconstruction by a substantial margin for hyperspectral imagery, but also to impose a much lighter computational burden. In essence, CPPCA uses $\mathbf{P}$ to project $\mathbf{X}$ within the sensing device, i.e., at the CPPCA sender.[1] At the receiver side of the system, the CPPCA reconstruction, given only random projections $\tilde{\mathbf{Y}}$ and

---

[1]Technically, CPPCA uses J distinct random projections $\left\{\mathbf{P}^{(j)}\right\}_{j=1}^{J}$ in order to enable the eigenvector recovery via projections onto convex sets (see [1]). For notational simplicity, we drop the j superscript on $\mathbf{P}$ as it does not impact the spectral-partitioning discussion of Section III, whereas [5] already addresses the issue for the spatial partitioning discussed in Section IV.

orthonormal matrix $\mathbf{P}$, recovers not only the PCA transform coefficients $\check{\mathbf{X}}$, but also approximations to the basis vectors (i.e., the eigenvectors) of the PCA transform $\boldsymbol{\Psi}$ for the dataset. Using these PCA eigenvectors, a simple linear least-squares pseudoinverse suffices to recover the PCA transform coefficients,

$$\check{\mathbf{X}} = \left(\mathbf{P}^{T}\boldsymbol{\Psi}\right)^{+}\tilde{\mathbf{Y}}. \tag{4}$$

The CPPCA receiver then simply estimates the original dataset as $\hat{\mathbf{X}} = \boldsymbol{\Psi}\check{\mathbf{X}}$.

We note that CS-based reconstruction relies on an assumption of sparsity in some transform domain, i.e., reconstruction with a union-of-subspaces paradigm [11] in which the signal is sparse in a basis yet the pattern of sparsity is unknown. On the contrary, the CPPCA model is that of the signal residing in single, albeit unknown, low-rank subspace, such that the CPPCA eigenvector-recovery procedure finds this subspace from the random projections. We argue that the union-of-subspaces model inherent to CS does not match the reality of most hyperspectral datasets—commonly, a single hyperspectral scene is considered to consist of a low-rank mixing of a relatively small number of spectral endmembers (e.g., [12], [13]). Such a mixing model is ideally suited to spectral decomposition with PCA; CPPCA inherits this advantage, while the proposed partitioning further increases the low-rank character of the data.

III. SPECTRAL PARTITIONING

PCA applied spectrally to a hyperspectral image is very effective when the spectral bands are highly correlated, permitting a close approximation of spectral signatures by retaining only a few principal components. CPPCA capitalizes on this same phenomenon, using random projections to recover principal eigenvectors and their corresponding PCA coefficients to reconstruct the dataset. Like PCA, CPPCA relies on a high degree of correlation among spectral bands to produce a reconstruction that closely approximates the original spectral signatures. However, it has been observed [6] that, commonly, neighboring bands in hyperspectral data are much more correlated than bands that are spectrally distant from one another.

To improve the performance of PCA in such a case of low correlation between spectrally distant bands, SPCA [6] partitions the dataset spectrally into a number of distinct spectral partitions, each of which contains a number of highly correlated bands. Subsequently, PCA is applied on each spectral partition independently. As a consequence, SPCA not only provides an exponential reduction in the computational burden required for the eigendecomposition central to PCA, but it also tends to improve the representational efficacy of PCA since each PCA decomposition is applied to only the highly correlated bands within each partition.

We adapt the spectral-partitioning paradigm of SPCA to the CPPCA framework by partitioning the dataset spectrally and applying CPPCA in each spectral partition independently. On the sensing side of the system, this is tantamount to the projection matrix $\mathbf{P}$ of the CPPCA sender having a block-diagonal structure with matrices $\mathbf{P}_g$ along the diagonal, where g is the index of the spectral partition. In effect, this $\mathbf{P}_g$ projects the N_g spectral bands of partition g into a K_g-dimensional subspace; thus,

$\mathbf{P}_g$ has size $K_g \times N_g$, where $K = \sum_g K_g$ and $N = \sum_g N_g$. That is,

$$\tilde{\mathbf{y}}_m^{(g)} = \mathbf{P}_g^T \mathbf{x}_m^{(g)}, \tag{5}$$

where $\mathbf{x}_m^{(g)}$ is an $N_g \times 1$ vector of the N_g spectral bands in partition g, and $\tilde{\mathbf{y}}_m^{(g)}$ has size $K_g \times 1$. On the receiver side of the system, CPPCA reconstruction is applied in partition g using $\tilde{\mathbf{Y}}^{(g)} = \begin{bmatrix} \tilde{\mathbf{y}}_1^{(g)} & \cdots & \tilde{\mathbf{y}}_M^{(g)} \end{bmatrix}$ and $\mathbf{P}_g$ independently of the other partitions. That is, in this spectrally-partitioned CPPCA, we recover an approximation to the principal eigenvectors of partition g along with the corresponding PCA coefficients; these are used to produce a reconstruction of the N_g bands of spectral partition g, i.e., $\mathbf{X}^{(g)} = \begin{bmatrix} \mathbf{x}_1^{(g)} & \cdots & \mathbf{x}_M^{(g)} \end{bmatrix}$. Concatenating the reconstructions $\mathbf{X}^{(g)}$ from the individual partitions produces the final reconstruction of the N-band dataset $\mathbf{X}$.

We assume that, given a specific hyperspectral sensor, the partitioning of the spectral bands, and hence N_g, is fixed and determined in advance by the typical correlation structure expected for the sensor. As a consequence, given an overall subrate K/N, we must select K_g for each partition such that the target subrate is met. The straightforward approach would be to set K_g such that the same subrate is used across all partitions, i.e., so that $K_g/N_g = K/N, \forall g$. However, we adopt a more sophisticated solution that permits K_g to reflect the fact that certain bands are anticipated to have a more complex signal content; thus, their partitions should receive a higher subrate relative to the others.

Specifically, we use a weighted sampling procedure based on the concept of loading factors that has been used previously in band prioritization [14], [15]. Loading factor $r_{n,l}$ associated with eigenvector $\mathbf{w}_n = \begin{bmatrix} w_{n,1} & \cdots & w_{n,N} \end{bmatrix}^T$ and eigenvalue λ_n is

$$r_{n,l} = w_{n,l} \sqrt{\lambda_n}, \qquad l = 1, 2, \ldots, N, \tag{6}$$

whereas the variance of spectral band l is

$$\rho_l = \sum_{n=1}^{N} r_{n,l}^2. \tag{7}$$

Given the target subrate, K/N, or, equivalently, a budget number of measurements for the entire dataset, K, the number of measurements for partition g is then

$$K_g = \frac{K \sum_{l \in \Omega_g} \rho_l}{\sum_{l=1}^{N} \rho_l}, \tag{8}$$

where Ω_g is the set of spectral bands in partition g.

In experimental results that follow, we use data from two hyperspectral sensors, AVIRIS and HYDICE. Using a correlation matrix typical to these sensors, sensor-specific spectral partitioning is as indicated in Table I—for both sensors, four spectral partitions have been used. For various subrates, the number of measurements used in each partition, K_g, again a sensor-specific value, is tabulated in Table II.

TABLE I
SPECTRAL BANDS IN EACH SPECTRAL PARTITION

Sensor	Spectral Partition			
	1	2	3	4
AVIRIS	4–36	39–106	116–152	167–221
HYDICE	1–56	57–102	103–133	134–191

TABLE II
NUMBER OF MEASUREMENTS, K_g, IN SPECTRAL PARTITION g FOR AVIRIS ($N = 193$) AND HYDICE ($N = 191$)

K_g	AVIRIS					HYDICE				
	Subrate, K/N					Subrate, K/N				
	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
K_1	5	11	17	22	28	7	16	25	30	40
K_2	8	20	32	41	46	6	12	18	25	31
K_3	3	3	4	6	10	3	4	7	9	13
K_4	3	5	5	8	14	3	6	8	11	12
K	19	39	58	77	97	19	38	57	76	96

TABLE III
EFFECTS OF SPECTRAL AND/OR SPATIAL PARTITIONING ON ECCENTRICITY

Dataset	Partitioning			
	none	spectral	spatial	spectral-spatial
Jasper	41.9	440.7	256.5	**997.7**
Moffett	44.4	592.0	304.1	**1001.7**
DC Mall	20.7	362.8	140.5	**757.7**

IV. SPECTRAL-SPATIAL PARTITIONING

Different from the spectral partitioning proposed above, [5] adapts CPPCA to exploit spatial classification. Specifically, in the spatial-partitioning CPPCA variant of [5], the receiver first classifies each projection vector $\tilde{\mathbf{y}}_m$ into one of several spatial partitions using either a supervised classifier or an unsupervised clustering algorithm. After spatial partitioning, CPPCA reconstruction is employed independently within each spatial partition. We note that, in this spatial-partitioning approach to CPPCA, partitioning occurs only at the reconstruction side of the system; i.e., the spatial partitions in [5] are formed only during reconstruction while the sender side of the system is identical to that of the original CPPCA. As a consequence, spatial partitioning as proposed in [5] is sensor-independent. This stands in contrast to the spectral partitioning proposed in Section III which is sensor-specific and present at both the sender and receiver sides of the CPPCA system.

Although several spatially partitioned CPPCA reconstructions are described in [5], for the work here, we adopt the technique called "unsupervised class-dependent CPPCA" which is conceptually the most simple. In this approach, the entire dataset is first reconstructed using conventional (i.e., non-partitioned) CPPCA; unsupervised clustering is applied to partition the dataset into C spatial classes; and finally CPPCA reconstruction is applied again, only this time it is applied in each spatial partition independently. Here, we use K-means for the spatial clustering. As in [5], the subrate for all spatial partitions is identical and equal to the overall subrate, K/N.

The spatial partitioning of [5] and the spectral partitioning of Section III are not mutually exclusive; it is straightforward to combine the two to effectuate a variant of CPPCA with spectral-spatial partitioning. In such a system, the sender applies

| Subrate | Spatial | | | | Spatial-spectral | | | |
| | Distance | | | | Distance | | | |
	Euclidean	City block	Cosine	Correlation	Euclidean	City block	Cosine	Correlation
0.1	24.70	24.50	21.61	23.08	25.00	24.39	23.97	24.06
0.15	29.04	28.90	26.40	26.41	30.52	30.46	29.90	29.71
0.2	31.10	30.40	27.31	29.70	32.60	32.16	32.12	32.29
0.3	33.85	33.53	31.66	31.85	36.33	36.06	35.68	35.89
0.4	35.71	35.69	34.70	33.81	37.79	37.64	37.43	37.19
0.5	37.17	36.54	35.35	35.45	39.46	39.36	39.13	38.90

random projection in each spectral partition using a sensor-specific spectral partitioning and number of measurements (i.e., Tables I and II). The receiver then reconstructs the dataset from the random projections using the following procedure:

1) Apply CPPCA reconstruction in each spectral partition independently, concatenating the reconstructed partitions to yield an initial CPPCA reconstruction of the dataset.
2) Apply K-means clustering on the initial reconstruction, yielding a spatial partitioning of the dataset.
3) Couple the spatial partitioning with the spectral partitioning, producing spectral-spatial partitions within the random projections.
4) Apply CPPCA reconstruction in each spectral-spatial partition independently, producing the final reconstruction of the dataset.

V. EXPERIMENTAL RESULTS

A. Reconstruction Results

We use two popular AVIRIS hyperspectral images, "Jasper Ridge" scene 1 and "Moffett" scene 1, both of which have 224 spectral bands and a spatial size of 512×512 pixels; after removal of noisy and water-absorption bands, we have $N = 193$ spectral bands. Spatial clustering uses $C = 5$ classes; this number of classes is estimated using the procedure based on the Bayesian information criterion (BIC) as described in [5]. We also test using a popular HYDICE dataset, "Washington DC-Mall." This image has 191 highly correlated spectral bands and a spatial size of 300×300 pixels. Using the same BIC approach to estimate the number of spatial classes, we choose $C = 6$.

We now examine the effectiveness of the various partitioning strategies proposed above; we use the implementation of CPPCA available from its website.[2] In [1], it is argued analytically that, fundamentally, CPPCA depends on the data distribution being sufficiently eccentric, meaning that the principal eigenvalues are widely separated from one another. Theoretical results in [1] suggest that the accuracy of CPPCA reconstruction is directly proportional to the eccentricity of dataset; thus, increasing data eccentricity can lead to improved performance. Table III presents the eccentricity of the various hyperspectral datasets using the various partitioning strategies considered above. Here, eccentricity is measured as the ratio of the first and second eigenvalues; in Table III, eccentricity is calculated in each partition (i.e., using the eigendecomposition for just that partition) and then averaged over all partitions.

[2]http://www.ece.msstate.edu/~fowler/CPPCA/

Table III also reports the global eccentricity as calculated over the entire, non-partitioned dataset. Table III indicates that, although both spatial and spectral partitioning when used alone increase the average eccentricity with respect to the non-partitioned dataset, the combined spectral-spatial partitioning results in the greatest degree of eccentricity.

We now explore the performance at reconstruction from random projections. In all cases, an orthonormalized Gaussian random matrix is used to project each pixel vector into a lower-dimensional space. Reconstruction quality is measured with a vector-based SNR in dB that is averaged over the dataset; i.e., for original vector $\mathbf{x}_m$ and its corresponding reconstruction $\widehat{\mathbf{x}}_m$,

$$\mathrm{SNR}(\mathbf{x}_m, \widehat{\mathbf{x}}_m) = 10 \log_{10} \frac{\mathrm{var}(\mathbf{x}_m)}{\mathrm{MSE}(\mathbf{x}_m, \widehat{\mathbf{x}}_m)}, \qquad (9)$$

where $\mathrm{var}(\mathbf{x}_m)$ is the variance of the components of vector $\mathbf{x}_m$, and the mean squared error (MSE) is

$$\mathrm{MSE}(\mathbf{x}_m, \widehat{\mathbf{x}}_m) = \frac{1}{N} \left\| \mathbf{x}_m - \widehat{\mathbf{x}}_m \right\|^2. \qquad (10)$$

We then average the vector-based SNR over all M vectors of the dataset. Alternatively, we measure the quality of a reconstructed hyperspectral dataset using an average spectral angle, where the spectral angle in degrees between the reconstructed hyperspectral pixel vector and its corresponding original vector is averaged over the dataset; i.e., $\bar{\xi} = \mathrm{mean}\,(\xi_m)$ where

$$\xi_m = \angle\,(\mathbf{x}_m - \widehat{\mathbf{x}}_m). \qquad (11)$$

In our spatial-clustering implementation, we use K-means with city-block (or ℓ_1) distance after comparing several similarity metrics: Euclidean, cosine, city-block, and correlation. The results are tabulated in Table IV, revealing that the city-block and Euclidean distances for K-means yield the best reconstruction for both CPPCA variants considered.

For the AVIRIS datasets, we see from Figs. 1 and 2 that CPPCA with spectral-spatial partitioning yields average SNR substantially higher than that of the original, non-partitioned CPPCA; additionally, spectral-spatial partitioning marginally outperforms spatial or spectral partitioning alone except at the lowest subrate. Fig. 3 indicates that similar gains are achieved for the HYDICE dataset for which spectral-spatial partitioning consistently outperforms the other partitioning strategies by typically 1–5 dB. Similar conclusions are reached by considering average spectral angle as tabulated in Table V. There, we see that spectral-spatial partitioning consistently yields lower average spectral angle than the other partitioning strategies. Overall, these results indicate that the combination of spatial
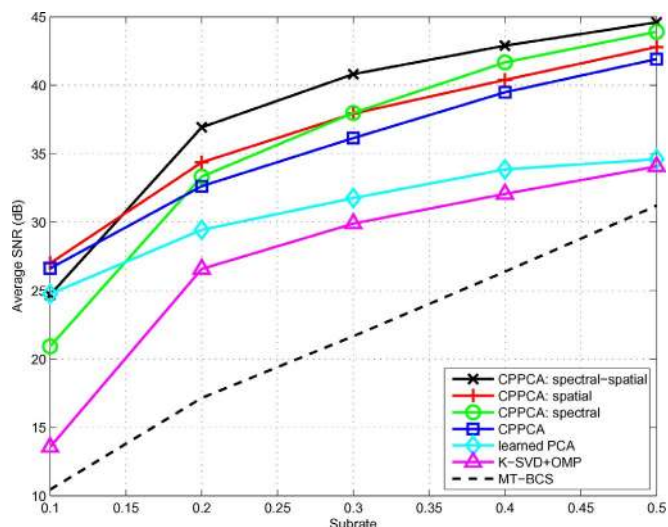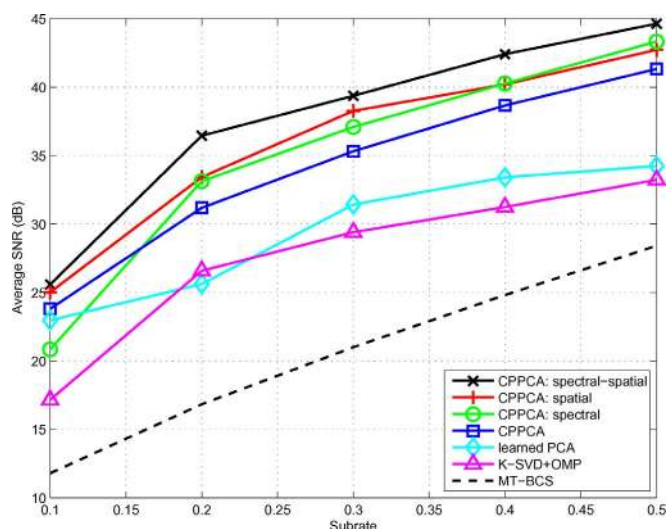
Fig. 1. Reconstruction for AVIRIS dataset "Jasper Ridge".

Fig. 2. Reconstruction for AVIRIS dataset "Moffett".

Fig. 3. Reconstruction for HYDICE dataset "Washington DC Mall". The curves for the CPPCA class are the results of averaging 100 trials using different random matrices $\mathbf{P}$. The corresponding standard deviations are illustrated with vertical bars.

TABLE V
AVERAGE SPECTRAL ANGLE $\bar{\xi}$ IN DEGREES

Algorithm	Subrate			
	0.2	0.3	0.4	0.5
Jasper				
CPPCA: spectral-spatial	**0.58**	**0.37**	**0.29**	**0.24**
CPPCA: spatial	0.84	0.55	0.41	0.3
CPPCA: spectral	0.92	0.52	0.34	0.26
CPPCA	1.01	0.66	0.44	0.33
learned PCA	1.36	1.04	0.82	0.75
K-SVD+OMP	4.67	1.96	1.4	1.16
MT-BCS	5.56	3.31	1.93	1.11
Moffett				
CPPCA: spectral-spatial	**0.57**	**0.41**	**0.28**	**0.21**
CPPCA: spatial	0.85	0.49	0.38	0.28
CPPCA: spectral	0.91	0.54	0.37	0.26
CPPCA	1.11	0.69	0.46	0.33
learned PCA	1.97	0.99	0.79	0.72
K-SVD+OMP	2.11	1.2	0.99	0.88
MT-BCS	5.37	3.32	2.14	1.41
DC Mall				
CPPCA: spectral-spatial	**0.99**	**0.65**	**0.55**	**0.45**
CPPCA: spatial	1.15	0.82	0.66	0.58
CPPCA: spectral	1.75	0.85	0.66	0.52
CPPCA	1.80	1.19	0.81	0.65
MT-BCS	4.97	3.14	2.29	1.77

and spectral partitioning yields the best reconstruction performance in terms of not only average SNR but also average spectral angle between pixels.

We also compare the partitioned variants of CPPCA to several competing CS-based reconstruction strategies. Figs. 1–3 and Table V present SNR and SAM results for MT-BCS [8] as described in Section II-A. Additionally, we compare also to K-SVD+OMP [9]. MT-BCS uses a Daubechies length-8 DWT as the fixed sparsity basis; on the other hand, K-SVD+OMP learns an overcomplete sparsity basis (size $= 193 \times 1296$) from training data. For this latter approach, we use the original "Moffett" dataset as training data when reconstructing the "Jasper Ridge" dataset, and vice-versa. Lacking a second HYDICE dataset that we can use for training, though, we provide K-SVD+OMP results for only the AVIRIS datasets (Figs. 1 and 2). We use implementations available from the respective authors of MT-BCS[3] and K-SVD+OMP.[4]

[3]http://people.ee.duke.edu/~lcarin/BCS.html. Note: due to excessive memory requirements of this implementation, the 512×512 datasets are split spatially into sixteen 128×128 partitions to which MT-BCS is applied independently.
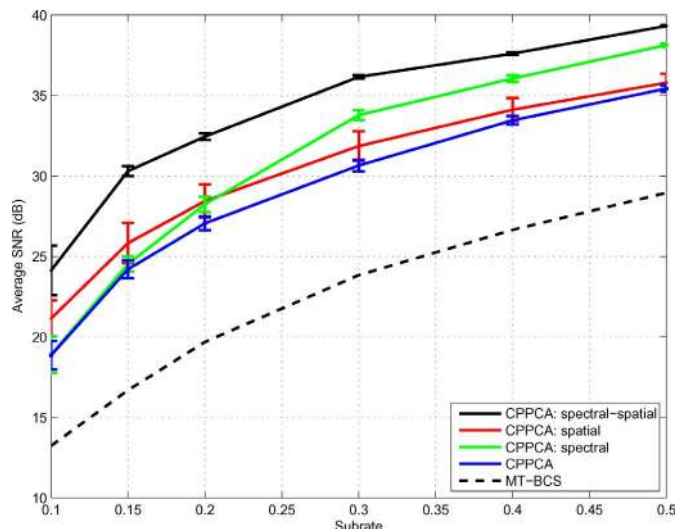
[4]http://www.cs.technion.ac.il/~elad/software/

Finally, as an alternative, albeit simplistic, "dictionary-learning" strategy, we propose to "learn" a dictionary of PCA eigenvectors by simply calculating the covariance matrix and its eigendecomposition for the training dataset, employing the resulting eigenvectors to yield a linear least-squares reconstruction of the randomly projected dataset. We see from Figs. 1 and 2 that, while K-SVD+OMP significantly outperforms the fixed-basis MT-BCS, the "learned-PCA" reconstruction does even better. However, all of the CPPCA-based reconstructions—which derive their PCA basis directly from the random projections of the dataset in question rather than from a separate body of training data—provide the best performance.

In order to examine the variability in reconstruction performance that is inherent due to the fact that the projection matrix $\mathbf{P}$ is randomly chosen, the SNR performance illustrated in Fig. 3 is

TABLE VI
RECONSTRUCTION TIME IN SECONDS FOR "MOFFETT," SUBRATE OF 0.3

Algorithm	Preliminary CPPCA Reconstruction	Spatial Clustering	Final CPPCA Reconstruction	Total
none	n/a	n/a	37	37
spatial	27	390	132	549
spectral	n/a	n/a	24	24
spectral-spatial	29	70	14	113
K-SVD+OMP†		2,372		
MT-BCS		67,261		

†time for OMP reconstruction only
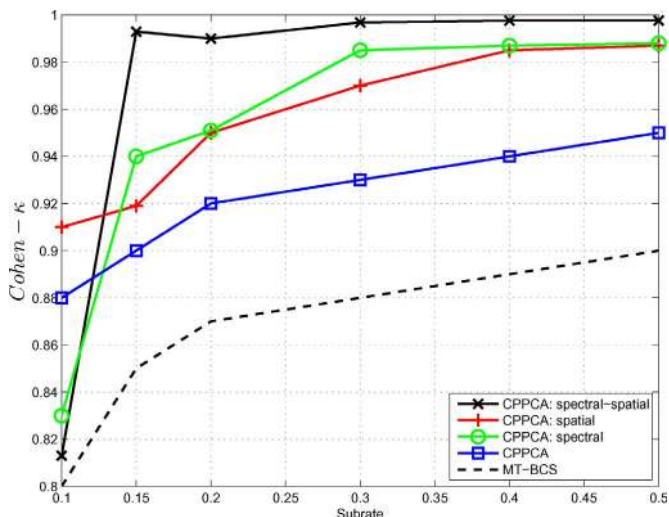
Fig. 4. Classification accuracy for HYDICE dataset "Washington DC Mall." using SVM.

Fig. 5. SAD (in radians) for "Cuprite" using VCA.

averaged over 100 trials, each using a different $\mathbf{P}$. Fig. 3 also indicates the standard deviation over the trials in the form of error bars.[5] We see that, not only does CPPCA with spectral-spatial partitioning yield higher average SNR over the trials, the variability in reconstruction quality due to the randomness of $\mathbf{P}$ is much smaller than that of the other techniques.

The execution time for the reconstruction algorithms under consideration is presented in Table VI. The results reveal that spectral partitioning is faster than the original non-partitioned CPPCA, while spectral-spatial partitioning is nearly five times as fast as spatial partitioning used alone. This speed up is as expected, since spectral clustering reduces the dimensionality of the CPPCA-reconstruction process, and CPPCA runs faster for smaller-dimensional datasets. Compared to other reconstruction algorithms, spectral-spatial CPPCA is nearly 20 and 600 times faster than K-SVD+OMP and MT-BCS reconstruction, respectively. We note also that the spectral partitioning and spectral-spatial partitioning could both be further expedited by implementing reconstruction in each spectral partition in parallel.

B. Supervised-Classification Results

We now investigate the class separability of the reconstructed data using supervised classification. Due to the availability of ground truth, we use the Washington DC Mall dataset, randomly choosing 10% of the labeled pixels in each class as training data while saving the remainder as testing data.

[5]Note that, due to the excessive run time of MT-BCS, Fig. 3 depicts the reconstruction from only one trial for MT-BCS.
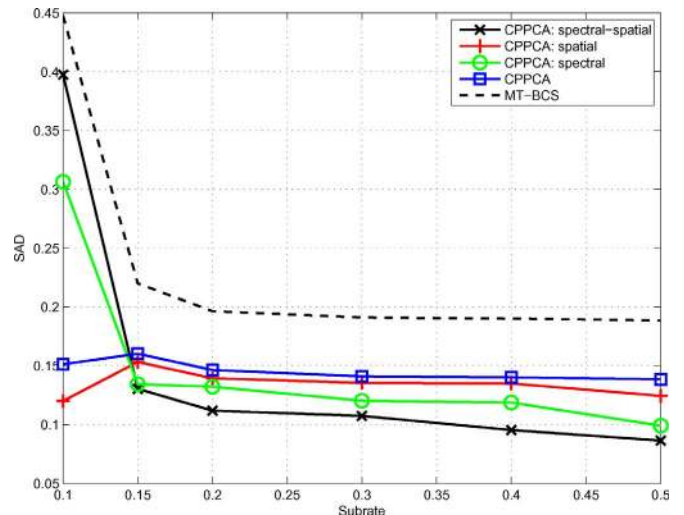
The popular kernel-based pattern-recognition algorithm, pixel-based support vector machine (SVM), is employed with a radial-basis (RBF) kernel. The SVM and RBF parameters are set to $(\varsigma, \gamma) = (8, 0.03125)$ via five-fold cross validation while the SVM-optimization problem is solved via the open-source LIBSVM package.[6] The accuracy of classification is evaluated using the Cohen-κ coefficient which is computed by weighting the measure accuracies and which shows a robust measure of the degree of agreement. We plot the classification accuracy, Cohen-κ coefficient, corresponding to various subrates in Fig. 4. As reported in Fig. 4, the spectral-spatial partitioned CPPCA consistently outperforms (higher κ value) the other partitioning strategies as well as MT-BCS except at the lowest subrate.

C. Linear-Unmixing Results

Vertex component analysis (VCA) [16] is one of the most popular methods for the linear unmixing of hyperspectral imagery. We apply VCA to evaluate the affect of the proposed method (spectral-spatial CPPCA) on unmixing. The Cuprite dataset with size of 250×181 pixels and 188 bands is considered in this section—water-absorption and noisy bands (including bands 1, 2, 104–113, 148–167, and 221–224) have been removed from the original 224-band dataset. We note that this same cropping of Cuprite was also used in [16] wherein the number of endmembers was set to $p = 14$. We apply various partitioning strategies coupled with CPPCA as well as MT-BCS, extracting p endmembers from the original as well as reconstructed images using VCA. Similar to [17], we calculate the average spectral-angle distance (SAD) between the endmembers obtained from the original image and those obtained from the reconstructed images. The results in Fig. 5 indicate that the proposed method (spectral-spatial CPPCA) gives better performance (smaller SAD) compared with other methods, except at the lowest subrate.

VI. CONCLUSIONS

In this paper, we have considered the use of both spectral and spatial partitioning in a system in which random projections are

[6]http://www.csie.ntu.edu.tw/~cjlin/libsvm

used at a sensor device to effectuate dimensionality reduction, while reconstruction takes place in a receiving device using CPPCA. Experimental results evaluated the quality of reconstruction not only in terms of SNR and spectral-angle fidelity, but also via performance of the reconstructed data in classification and linear-unmixing tasks. The results demonstrated that, although CPPCA with either spectral or spatial partitioning achieved significantly higher reconstruction quality than the original, non-partitioned CPPCA, the combination of the two outperformed either form of partitioning used alone and is far better than CS reconstruction using either fixed or learned bases. Additionally, the introduction of spectral partitioning was seen to significantly reduce the computational complexity of CPPCA reconstruction due to a much lower vector dimensionality in each spectral partition.

REFERENCES

[1] J. E. Fowler, "Compressive-projection principal component analysis," *IEEE Trans. Image Process.*, vol. 18, no. 10, pp. 2230–2242, Oct. 2009.

[2] J. E. Fowler and Q. Du, "Anomaly detection and reconstruction from random projections," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 184–195, Jan. 2012.

[3] A. Majumdar and R. K. Ward, "Robust classifiers for data reduced via random projections," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 5, pp. 1359–1371, Oct. 2010.

[4] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[5] W. Li, S. Prasad, and J. E. Fowler, "Classification and reconstruction from random projections for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, 2012, to be published.

[6] X. Jia and J. A. Richards, "Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 538–542, Jan. 1999.

[7] Q. Du, W. Zhu, H. Yang, and J. E. Fowler, "Segmented principal component analysis for parallel compression of hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 713–717, Oct. 2009.

[8] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 92–106, Jan. 2009.

[9] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[10] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[11] Y. M. Lu and M. N. Do, "Sampling signals from a union of subspaces," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 41–47, Mar. 2008.

[12] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[13] 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 J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[14] C.-I. Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, "A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 6, pp. 2631–2641, Nov. 1999.

[15] T.-M. Tu, C.-H. Chen, J.-L. Wu, and C.-I. Chang, "A fast two-stage classification method for high-dimensional remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 1, pp. 182–191, Jan. 1998.

[16] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 989–910, Apr. 2005.

[17] F. García-Vílchez, J. M. noz Marí, M. Zortea, I. Blanes, V. González-Ruiz, G. Camps-Valls, A. Plaza, and J. Serra-Sagristà, "On the impact of lossy compression on hyperspectral image classification and unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 253–257, Mar. 2011.
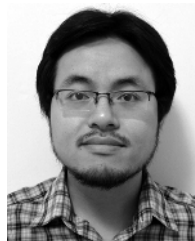
Nam Hoai Ly (S'11) received the B.E. degree in electronics and telecommunications from Hanoi University of Science and Technology, Hanoi, Vietnam, in 2008. He is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS.

Since 2009, he has been a Research Assistant with the Geosystems Research Institute, and a Teaching Assistant with the Department of Electrical and Computer Engineering, Mississippi State University. His research interests include hyperspectral image processing, machine learning, and computer vision.

Qian Du (S'98–M'00–SM'05) received the Ph.D. degree in electrical engineering from the University of Maryland Baltimore County in 2000.

She was with the Department of Electrical Engineering and Computer Science, Texas A&M University–Kingsville, from 2000 to 2004. She joined the Department of Electrical and Computer Engineering at Mississippi State University in Fall 2004, where she is currently an Associate Professor. Her research interests include hyperspectral remote sensing image analysis, pattern classification, data compression, and neural networks.

Dr. Du currently serves as Co-Chair for the Data Fusion Technical Committee of IEEE Geoscience and Remote Sensing Society. She also serves as Associate Editor for IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS) and Associate Editor for IEEE *Signal Processing Letters*. She received the 2010 Best Reviewer award from IEEE Geoscience and Remote Sensing Society. She is the General Chair for the 4th IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) in Shanghai, China in 2012. Dr. Du is a member of SPIE, ASPRS, and ASEE.

James E. Fowler (S'91–M'96–SM'02) received the B.S. degree in computer and information science engineering and the M.S. and Ph.D. degrees in electrical engineering in 1990, 1992, and 1996, respectively, all from the Ohio State University.

In 1995, he was an intern researcher at AT&T Labs, Holmdel, NJ, and in 1997, he held an NSF-sponsored postdoctoral assignment at the Université de Nice-Sophia Antipolis, France. In 2004, he was a visiting Professor in the Département TSI at École Nationale Supérieure des Télécommunications (ENST), Paris, France. He is currently Billie J. Ball Professor and Graduate Program Director of the Department of Electrical and Computer Engineering at Mississippi State University, Starkville, MS. He is also a researcher in the Geosystems Research Institute (GRI) at Mississippi State.

Dr. Fowler is an Associate Editor for IEEE TRANSACTIONS ON IMAGE PROCESSING and the EURASIP *Journal on Image and Video Processing*. He previously served as an Associate Editor for IEEE TRANSACTIONS ON MULTIMEDIA and IEEE *Signal Processing Letters*. He is the chair of the Image, Video, and Multidimensional Signal Processing Technical Committee of the IEEE Signal Processing Society and the publicity chair of the program committee for the Data Compression Conference.