CLASSIFICATION PERFORMANCE OF RANDOM-PROJECTION-BASED DIMENSIONALITY REDUCTION OF HYPERSPECTRAL IMAGERY

James E. Fowler[‡], Qian Du, Wei Zhu, and Nicolas H. Younan
Department of Electrical and Computer Engineering, Geosystems Research Institute,
Mississippi State University, USA
[‡]corresponding and presenting author, fowler@ece.msstate.edu

In the traditional data pipeline used with high-dimensional data such as hyperspectral imagery, data is acquired in its full dimensionality within some typically remote signal-acquisition platform (e.g., a satellite) only to be often subsequently reduced in dimension prior to application-specific processing such as classification. However, it would be greatly beneficial if dimensionality reduction could occur before data downlink, since many signal-acquisition platforms are severely resource-constrained such that on-board dimensionality reduction could dramatically cut storage and communication burdens faced by such remote sensors. However, many approaches to dimensionality reduction are not only data dependent but also exceedingly computationally expensive so as to preclude on-board implementation. In this work, we investigate a significant departure from the traditional data flow by integrating dimensionality reduction directly into the signal-acquisition process in order to avoid not only the computational burden of explicit dimensionality reduction, but also the production of onerous quantities of data in the first place. We employ the emerging mathematical theory of random projections to this end, effectively shifting the computational burden of dimensionality reduction from the resource-constrained acquisition device to a more powerful base-station system which is presumably an earth-based central site.

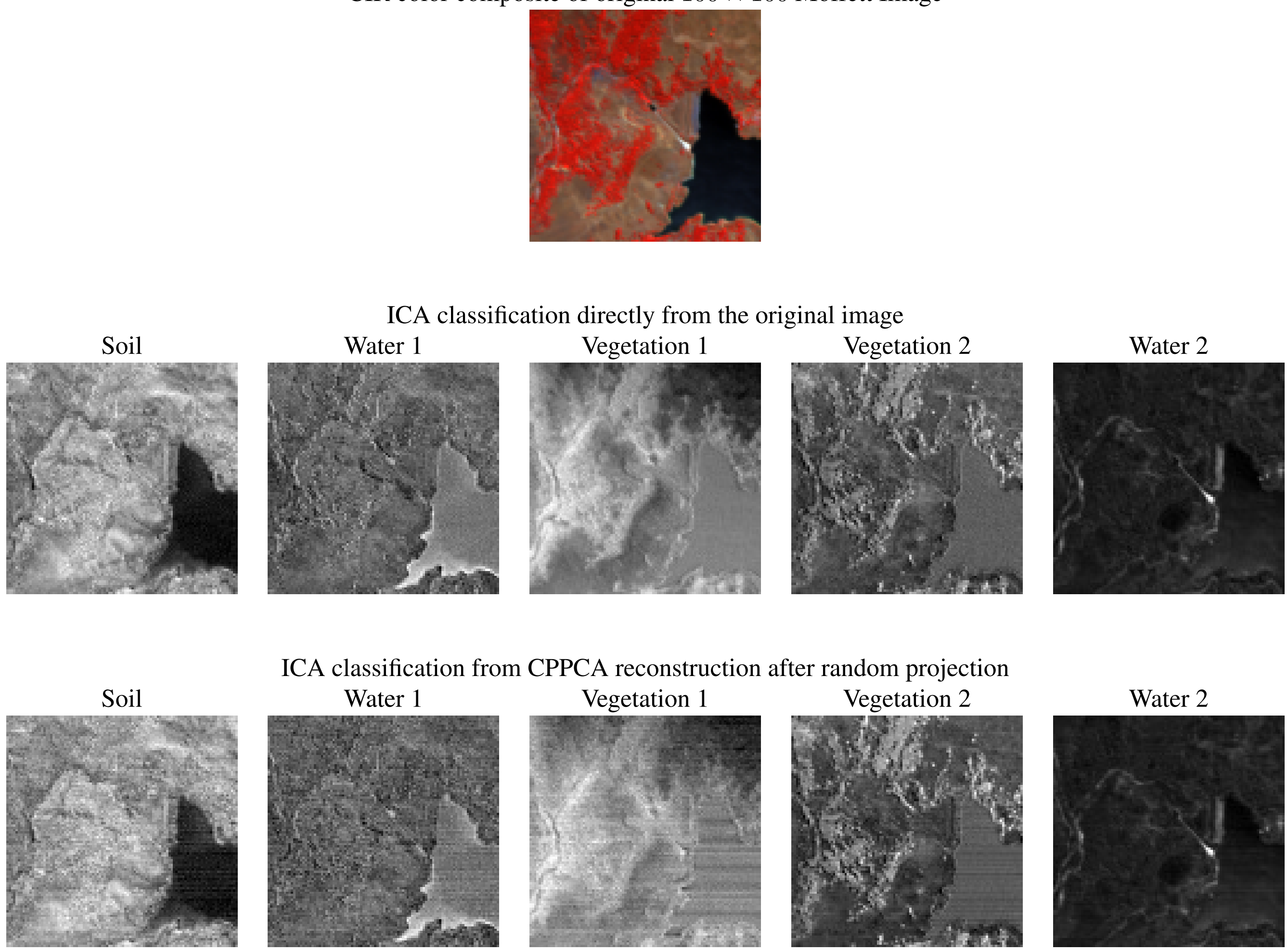
Recent work in the applied-mathematics and signal-processing communities has demonstrated that projections onto randomly chosen subspaces can be a particularly powerful form of dimensionality reduction. Namely, the mathematical theory of *compressed sensing* (CS) (e.g., [1–4]) establishes that sparsely representable signals can be recovered exactly from data-independent random projections. Furthermore, we have recently developed *compressive-projection principal component analysis* (CPPCA) [5, 6], which recovers an approximate PCA representation of the original signal from random projections. Both CS and CPPCA permit sensing platforms to enjoy the benefits of dimensionality reduction (less burdensome storage and communication requirements) without the expense of computation associated with explicit dimensionality reduction since the random projections can be accomplished simultaneously with the sensing and signal-acquisition process, while the more expensive reconstruction from the projections takes place at the receiver-side base station.

The merits of random projection plus CS or CPPCA reconstruction have previously been established mathematically with a focus on squared-error-based signal-quality measures. On the contrary, the main objective of the present work focuses on performance at classification and the corresponding issue of preservation of statistical class separability for random-projection-based dimensionality reduction, since such application-specific classification performance is typically of primary importance in many geospatial applications.

Previous results in [5, 6] indicate that CPPCA outperforms CS for a variety of bases when reconstruction quality is measured in terms of a traditional squared-error-based measure. We anticipate that this is due in no small part to the fact that CPPCA recovers an approximate PCA basis for the dataset in question, and PCA is known to be the optimal basis in several mathematical respects in terms of reduction of squared error per percentage of coefficients retained. It is known, however, that PCA is not necessarily the optimal basis for maintaining class separability.

Fig. 1 presents preliminary experimental results regarding the classification performance resulting from CPPCA reconstruction for a 100×100 spatial region of the "Moffett" hyperspectral image, an AVIRIS image with 224 spectral bands. The second row of images in Fig. 1 depicts the five independent components determined from the original image (shown on the first row of Fig. 1); these component images are produced using the Joint Approximate Diagonalization of Eigenmatrices (JADE) algorithm for ICA. The third row of images in Fig. 1 depicts the same independent component images determined from a CPPCA reconstruction after the original image was reduced in spectral dimension by random projection onto 90-dimensional random subspaces (i.e., a 40% reduction in data size). Also included is an objective measure of similarity, the spatial correlation coefficient as denoted by ρ. Since, for each of the endmember classes, ρ is quite close to 1.0, we conclude that the random projections are successful in preserving essentially all of the class separability in this example.

CIR color composite of original 100×100 Moffett Image



$\rho = 0.9771$ $\rho = 0.9137$ $\rho = 0.9360$ $\rho = 0.9528$ $\rho = 0.9867$

Figure 1: Classification performance for CPPCA as compared to that on the original image

1. REFERENCES

[1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, February 2006.

[2] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, December 2006.

[3] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

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

[5] J. E. Fowler, "Compressive-projection principal component analysis for the compression of hyperspectral signatures," in *Proceedings of the IEEE Data Compression Conference*, J. A. Storer and M. W. Marcellin, Eds., Snowbird, UT, March 2008, pp. 83–92.

[6] ——, "Compressive-projection principal component analysis and the first eigenvector," in *Proceedings of the IEEE Data Compression Conference*, J. A. Storer and M. W. Marcellin, Eds., Snowbird, UT, March 2009, pp. 223–232.