SPATIALLY-COHERENT NON-LINEAR DIMENSIONALITY REDUCTION AND SEGMENTATION OF HYPER-SPECTRAL IMAGES

By

Anish Mohan

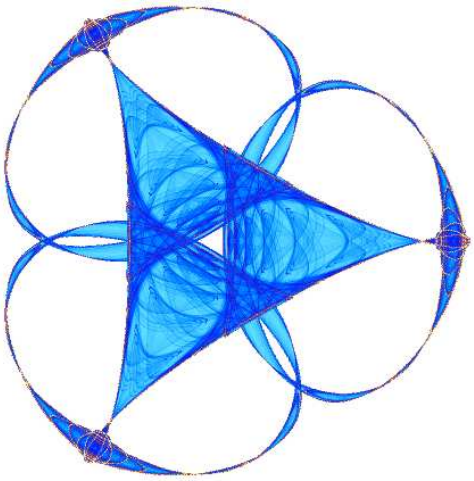
Guillermo Sapiro

and

Edward Bosch

IMA Preprint Series # 2123

(June 2006)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

400 Lind Hall

207 Church Street S.E.

Minneapolis, Minnesota 55455–0436

Phone: 612/624-6066 Fax: 612/626-7370

URL: http://www.ima.umn.edu

Spatially-Coherent Non-Linear Dimensionality Reduction and Segmentation of Hyper-Spectral Images

Anish Mohan,[1] **Guillermo Sapiro,**[1] and **Edward Bosch**[2]

1. Electrical and Computer Engineering and IMA, University of Minnesota, Minneapolis, MN 55455, {anish,guille}@ece.umn.edu

2. National Geospatial-Intelligence Agency, 12310 Sunrise Valley Drive, Reston, VA 20191, Edward.H.Bosch@nga.mil

***Abstract*— Non-linear dimensionality reduction and vector segmentation of hyper-spectral images is investigated in this letter. The proposed framework takes into account the non-linear nature of high dimensional hyper-spectral images, and projects onto a lower dimensional space via a spatially-coherent locally linear embedding technique. The spatial coherence is introduced by comparing individual pixels based on their local surrounding neighborhood structure. This neighborhood concept is also extended to the segmentation and classification stages using a modified vector angle distance. We present the underlying concepts of the proposed framework and experimental results showing the significant classification improvements.**

I. Introduction

Hyper-spectral images are produced by sensors such as AVIRIS (Airborne Visible/Infrared Imaging Spectrometer, a NASA/Jet Propulsion Laboratory Sensor) and Hyperion (a NASA sensor). These sensors provide data in the form of hundreds of narrow and adjacent spectral bands. This spectrum is often unique to the material composition and could be used, for example, to investigate the compounds and elements on the surface of the Earth. Such identification is of great significance for detecting minerals associated with ore deposits [4], precision farming by assessing crop and soil conditions [18], and inland water environmental monitoring [7], to name just a few of the numerous applications of hyper-spectral data.

Hyper-spectral sensors such as AVIRIS deliver calibrated images of 224 contiguous spectral channels (bands), with wavelengths from 400 to 2500 nm [1]. The amount of data that is acquired may typically run into hundreds of megabytes. The processing of this amount of data requires considerable time and computing resources. The work here presented first addresses the problem of reducing this huge amount of data to tractable levels. The redundancy in the data of adjacent bands is intrinsically exploited to help in this data reduction. Simultaneously, as shown with the experiments in this work, this data reduction also removes spurious and erroneous information from the data, thereby leading to more accurate clustering and classification. We thereby introduce a framework for dimensionality reduction of hyper-spectral images by adapting the theory of non-linear dimensionality reduction. Specifically, our efforts are directed to extend and adapt locally linear embedding techniques, [17], to the dimensionality reduction of hyper-spectral images.

Early research in dimensionality reduction and classification of hyper-spectral images focused on linear projection methods such as Principal Component Analysis (PCA) [16]. Methods like PCA, Factor analysis, and multidimensional scaling, assume that the underlying data manifold is linear, which is not necessarily true in the case of hyper-spectral data and most imaging modalities. PCA has been shown not to perform optimally in the presence of interference sources such as natural background signatures and structured non-random noise such as striping, [6], and it is also not appropriate for material identification and separability. Orthogonal subspace projection, introduced in [11], does not exploit the higher order correlations between the spectral bands nor it addresses the non-linear mixture of spectral signature as explained in [14], which presents a kernel based non-linear version to address these issues. In linear spectral mixing methods, [12], the basic idea is that each pixel in the image can be decomposed into its constituent end members. This approach does not address the issue of non-linear mixing of spectral signatures. The idea of selecting the best bands for analyzing the hyper-spectral data was also investigated, [9]. The thrust behind this approach was that each object has unique spectral features which can be identified by looking at particular bands.

Most of these established methods do not consider the non-linear characteristics of the hyper-spectral data. The multiple sources of nonlinearity have been pointed out in [2], and include the nonlinear nature of scattering, non-linearities due to the variable presence of water in pixels as a function of position in the landscape, multiple scattering within a pixel, and the heterogeneity of sub pixel constituents

A number of methods have been proposed in the literature in the domain of non-linear dimensionality reduction and manifold learning. Data driven and non-linear dimensionality reductions algorithms, e.g., [3], [17], [19], have become popular in the last few years and are the critical mainstream in high dimensional data analysis. These methods try to find the underlying structure of the sampled manifold by non-linear projections of sample data points. Although these methods are not strictly based on any physical or phenomenological models, nevertheless they provide a powerful framework for processing high dimensional data. Applications of non-linear dimensionality reduction methods such as ISOMAP in hyper-

spectral images is explained in [2] (the issue of tiling for large images is also addressed in this work). Unfortunately, ISOMAP has been shown to work only under quite limited conditions [8]. Instead, we propose to use the method of *locally linear embedding* (LLE) [17]. Early uses of LLE for hyper-spectral data were reported in [13]. Here, we first extend LLE by introducing spatial coherence, which is here shown to be critical for accurate classification and segmentation. This spatial coherence is introduced following recent results in the literature of image denoising and texture synthesis [5], [10], [15], where pixels are compared based on their local $n \times n$ immediate surrounding neighbors. The same concept of spatial coherence is then also introduced into vector angle distances in order to classify and segment the data from the non-linearly embedded hyper-spectral images.

After presenting the original LLE technique, Section II, we describe the proposed spatially-coherence extension, Section III, as well as its introduction into the vector angles distance measurement. We then experimentally show the advantages of the proposed framework, Section IV. Concluding remarks are then presented in Section V

II. Locally Linear Embedding

Locally linear embedding, [17], is a data driven non-linear dimensionality reduction method. It is based on the assumption that each sample point in the higher dimensional manifold can be approximated by a linear combination of its local neighbors. It is further assumed that the neighborhood relationship is preserved during the process of dimensionality reduction, meaning that for the given high dimensional data with a low intrinsic (co-)dimension, the same neighborhoods exist in the original high and lower projected dimensions. The local geometry for each data point is characterized by patches formed by the coefficients that linearly reconstruct the data point from its neighbors. For every data point $\vec{X_i} \in R^D$ in high D dimensions, the weights $\mathcal{W} = \{W_{ij}\}_{i,j}$ are computed by minimizing the cost function

$$\epsilon(\mathcal{W}) := \sum_i \|\vec{X_i} - \sum_{j \in N(i)} W_{ij}\, \vec{X}_{ij}\|^2, \qquad (1)$$

where the external sum is over all the data points $\vec{X_i}$, and the internal sum is over $N(i)$, the k closest neighbors of $\vec{X_i}$ (see Section III). The unknown weights W_{ij} are constrained to add to one ($\sum_{j \in N(i)} W_{ij} = 1$).

After the weights are computed, each high dimensional observation is mapped to a low dimensional space preserving the local structure of the manifold. This is done by choosing low dimensional coordinates $\vec{Y_i} \in R^d$, $d \ll D$, minimizing the embedding cost function ($\mathcal{Y} = \{\vec{Y_i}\}_i$)

$$\phi(\mathcal{Y}) := \sum_i \|\vec{Y_i} - \sum_{j \in N(i)} W_{ij}\, \vec{Y}_{ij}\|^2 \,. \qquad (2)$$

Note that while in (1) the minimization is performed on the weights W_{ij}, these are then fixed in (2), and the minimization is done with respect to the coordinates $\vec{Y_i}$, for which a lower dimensionality has been selected. The minimizations are straightforward applications of singular valued decomposition [17].

III. Spatially-Coherent Neighborhood Selection

The original LLE algorithm has no implicit spatial coherence. Each data is considered a point sample in the underlying high dimensional manifold embedded in R^D. In the case of hyper-spectral images, these data points are pixels. The pixel-vector, $\vec{X_i} \in R^D$ in Equation (1), is then the vector containing the pixel value at a particular position for all the bands. Hyper-spectral data are images, and hence the pixel-vectors are also spatially related. For example, if a given pixel-vector belongs to a particular class, the probability of its immediate surrounding pixel-vectors belonging to the same class is high. Our aim is to first introduce this spatial coherence in the computation of the high dimensional local neighborhood $N(i)$ in equations (1) and (2).

In classical locally linear embedding, the Euclidean distance is used to compute $N(i)$, the k closest local neighbors to a point $\vec{X_i}$. Instead of using the individual pixels of each band to compute this Euclidean distance, we propose to use all the $n \times n$ surrounding pixels to $\vec{X_i}$ (an $n \times n$ square with $\vec{X_i}$ as its center). Hence, each coordinate (band) in $\vec{X_i}$ is replaced by an n^2-dimensional vector (we use $n = 3$ in this work). In other words, in the case of Euclidean distances, instead of computing the distance between $\vec{X_i}$ and $\vec{X_j}$ as

$$d_E(\vec{X_i}, \vec{X_j}) := \sqrt{\sum_p d_e^2(\vec{X_i}(p), \vec{X_j}(p))}, \qquad (3)$$

where $d_e(\vec{X_i}(p), \vec{X_j}(p)) := |\vec{X_i}(p) - \vec{X_j}(p)|$, p runs over the dimensions of $\vec{X_i}$, and $\vec{X_i}(p)$ stands for the respective coordinate (band); we use the following new spatially-coherence distance to compute the $\vec{X_i}$ neighborhood $N(i)$:

$$d_S(\vec{X_i}, \vec{X_j}) := \sqrt{\sum_p d_{\mathcal{N}}^2(\vec{X_i}(p), \vec{X_j}(p))}, \qquad (4)$$

where now $d_{\mathcal{N}}(\vec{X_i}(p), \vec{X_j}(p))$ is the distance between the corresponding $n \times n$ surrounding pixels to $\vec{X_i}(p)$ and $\vec{X_j}(p)$ (we can use for example Euclidean distance, L_1 distance, or vector angles to compare these two pixel patches for example). We have then replaced in the original Euclidean distance the pixel wise differences by $n \times n$ patches differences, for each coordinate p, thereby introducing spatial coherence. This substitution has been recently found to be crucial for state-of-the-art image denoising as well [5], [15].

Note that we do not have to restrict our attention to d_S in Equation (4) being Euclidean, and we can also consider here L_1 or vector angle distances for example. We experiment with these variations as well.

Following this definition of distance, if two pixels are very similar but their respective surrounding $n \times n$ patches are dissimilar, then the distance d_S between the two pixel-vectors

would be high and such pixels would not be considered neighbors and will not be jointly used to estimate the local linear embedding in equations (1) and (2). The distance d_S between the two vectors inherently takes into account the spatial coherence between two vector-pixels. Comparing the neighborhoods as a measure of similarity for pixel-vectors provides robustness to spurious information for a pixel, e.g., noise.

A different approach to calculate the distance between pixel-vectors is to explicitly introduce a distance matrix $\mathcal{M}$ whose entries are proportional to the spatial distance between pixels. This new spatial-distance matrix can then be used to weigh the entries of the final distance matrix between the pixel-vectors. A problem with this is that if objects of the same class are spatially separated, then this spatial-distance matrix would weigh down the contribution of the spatially separated, although of same class, pixels, which is undesirable.

Finally, we should note that this concept of replacing individual pixel differences by $n \times n$ surrounding block differences is quite general, and as such, we use it below also to compute the vector angle difference between two pixel-vectors for classification and clustering.

IV. Experimental Results

Experimental results are now presented for two different data sets. The first data set is the hyper-spectral image of the Indian Pines, obtained by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS, a NASA/JPL sensor) in 1992.[1] This data set has 224 spectral channels covering a spectral region from 0.4 to 2.5 mm in 10 nm intervals. Out of the 224 bands, about 43 bands contained unstructured noise. These bands are excluded from the experiments, resulting in 181 bands.[2] This data comes with its corresponding ground truth classification, and is thereby here used to show the importance of spatially-coherent non-linear dimensionality reduction and classification. The second data is obtained from the Hyperion sensor aboard the EO-1 satellite (a NASA sensor). This data set contains 242 spectral channels covering a spectral region from 356nm to 2577nm, and has a spatial resolution of 30 meters. Out of the 242 bands, 44 bands are not calibrated due to decreased sensitivity of the detectors within the non-calibrated spectral regions. The non calibrated bands are set to zero, and hence we exclude these bands in our experiments. The results shown in this paper are for two random image segments of 30×30 and 36×36 pixels, for the AVIRIS satellite, and two segments of size 70×70 and 80×80 pixels for the image obtained from the EO-1 satellite. Similar results where obtained for other regions of the hyper-spectral images.

For the AVIRIS data, the original 181 band image was first reduced using the here proposed spatially-coherent (modified) locally linear embedding, projection onto 5, 10, 25, 50, and 100 bands. For each of these experiments we varied the number of neighbors (cardinality of $N(i)$), using 5, 10, 25, 50, 100, 200, 300, and 400 neighbors. Once the dimension reduced (projected) image was obtained, we performed the classification using the spatially-coherent vector angles measure. For the first image segment, which contains four classes, a sample of size 50 pixels/per class were used as a reference vectors. For the second image segment, which contains 8 different classes, a sample of 18 pixels/per class were taken as the reference vectors. Both images also had unclassified background segments which did not have a representative spectra, and background segmentation is not addressed in this test. To overcome this, we forced each pixel in the image to take a label of one of the existing 4/8 classes. The background pixels are excluded while considering the accuracy of the classification.

In the following images from AVIRIS, for finding the neighbors $N(i)$ to estimate the locally linear embedding, we experimented with three different distance metrics for comparing the vector-neighborhoods (both for d_S and for $d_{\mathcal{N}}$): Euclidean, vector angle, and L_1. For simplicity of the presentation, we didn't consider all the freedom of the model in Equation (4), and test only for d_S and $d_{\mathcal{N}}$ being both the same selection from the three metrics mentioned above.[3] We also experimented by introducing the spatial distance matrix $\mathcal{M}$ in conjunction with the distance matrix $\{d_S(\vec{X}_i, \vec{X}_j)\}_{i,j}$ obtained from the above three distance metrics, see Equation (4). For the first image from AVIRIS, the classification accuracy obtained by using all the original 181 noise-free bands was about 87.8%. With our modified locally linear embedding, we were able to reduce the number of bands to 10 and obtain a classification accuracy of 99.82%. Reducing the 181 band to 10-25 bands gave consistent accuracy of more than 90% across all tested distance metrics d_S and the number of neighbors considered. Reduction to 10 bands considering 5 neighbors (according to the spatially consistent metric defined above), gave the best result consistently for all 3 distance metrics. Reduction to 25 dimensions considering 10 neighbors gave the best classification results. Weighing by the spatial-distance matrix $\mathcal{M}$ consistently improved the results for all three distant metrics and for all the considered number of projected bands. The results for reduction onto 10 and 25 bands were comparable. The classification accuracy decreased when the reduced number of bands were 5 (too few) or 50 (too many). Results for this image are visualized in Figure 1.

For the second image from AVIRIS, the classification accuracy obtained by using all the original 181 noise-free bands was about 67%. With our modified, spatially coherent, locally linear embedding we were able to reduce the bands to 25 and obtain a classification accuracy of 85%. Projecting the 181 bands to a 25 band image gave consistent accuracy of more than 75% across all tested distance metrics and the number of neighbors considered. Reducing to 25 bands and

[1] http://www.ece.purdue.edu/∼biehl/MultiSpec/documentation.html

[2] This task could be easily automated by using the fact that the noisy bands lack structure.

[3] In other words, we use Euclidean, L_1, or vector angle to compare vectors in $R^{9 \times 181}$, the 9 coming from the 3×3 neighborhood and the 181 from the number of bands. We could obtain even better results if $d_{\mathcal{N}}$ uses one distance, e.g., vector angle, and then d_S a different one, e.g., Euclidean as in Equation (4).

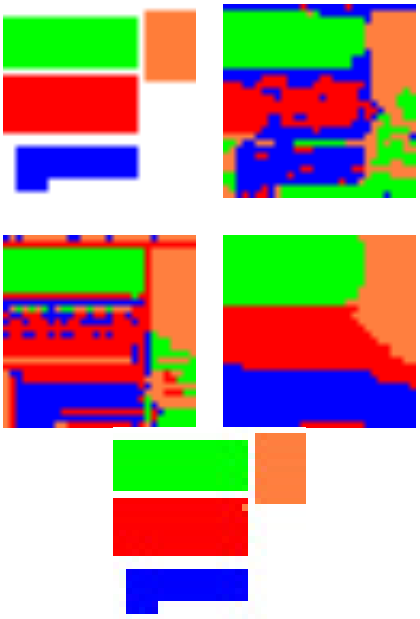


Fig. 1. *Top left:* The Ground truth data. Each of the four classes is represented by different colors. White corresponds to background. *Top right:* Classification using the original 181 noise free bands. The classification accuracy is 87.8%. *Middle left:* Classification after reducing the 181 band image to 10 bands using LLE. Vector Angle was used for classification. The classification accuracy is 86.91%. *Middle right:* Classification after reducing the 181 band image to 10 bands using the proposed spatially-coherent LLE combined with the spatially-coherent vector angle for classification. The classification accuracy is 99.82 %. *Bottom:* Same as on the middle right, with the background pixels masked. Compare with ground truth on top left.

considering 5 neighbors gave the best result consistently for all 3 distance metrics. The results when projecting onto 10 bands are marginally lower. The classification accuracy decreases by 5-10% when the reduced number of bands were 5 or 50. Weighing by the spatial-distance matrix decreases the accuracy of classification. This is due to the fact that in this image, there is one class which is spatially separated. Results for this data are presented in Figure 2.

A confusion matrix for this data is presented in Table I. The label for each row shows the ground truth class, and labels in the columns show the classified class. This table shows for each class how many vector-pixels have been correctly classified and how many have been not. In the miss-classification case, the class they have been assigned to is shown. As expected, most of the vector-pixels are correctly classified, as can be seen from the large numbers in the diagonal and the otherwise sparsity of the matrix. The spectral signature of the ground truth BLU and the ORN region are very close to each other, see Figure 3, therefore we get maximum miss-classification when BLU region falsely identified as ORN region.

For the second set of images obtained from the EO-1 satellite (two segments of it), the original data set of 242 bands was reduced to 172 bands and 178 bands respectively by removing the non-calibrated bands. Though 42 bands are

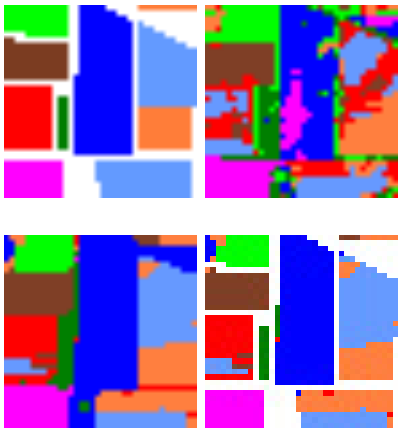


Fig. 2. *Top left:* The ground truth data. Each of eight classes is represented by different colors. White represents background. *Top right:* Classification using the original 181 noise free bands. Background pixels are classified as belonging to one of existing eight classes. The classification accuracy is 67%. *Bottom left:* Classification after reducing the 181 band image to 10 bands using our modified LLE and the modified vector angle for classification. The classification accuracy goes to 85%. *Bottom right:* Same as on the left, with the background pixels masked. Compare with ground truth on top left.

	ORN	RED	GRN	MGT	DBL	DGR	BLU	BRN
ORN	72	0	0	0	0	0	15	4
RED	0	83	0	0	0	0	18	7
GRN	6	0	54	0	8	0	0	0
MGT	0	0	0	77	0	0	0	0
DBU	0	1	0	0	279	6	0	0
DGR	0	0	0	0	0	20	0	0
BLU	80	3	0	0	2	0	184	0
BRN	1	0	0	0	0	0	0	85

TABLE I

CONFUSION MATRIX

The rows of the confusion matrix represent the true label of each class, as obtained from the ground truth data. Columns represent the class assigned by our proposed algorithm. Color codes are as follows ORN-Orange: RED-Red: GRN-Green: MGT-Magenta: DBU-Dark Blue: DGR-Dark Green: BLU-Blue: BRN-Brown

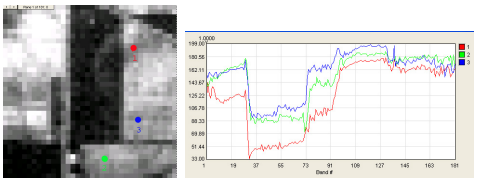


Fig. 3. Explanation for the BLU and ORN miss-classification shown in the confusion matrix. Note, as classified by the ground truth, vector-pixels from the same class (green and red) are further away than vector-pixels from different classes (green and blue).

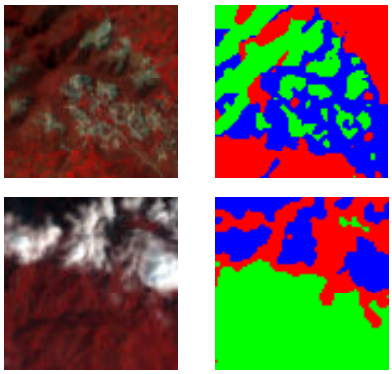


Fig. 4. *Top left:* Image from E0-1 satellite. The three visible classes are mountain tops, plains, and green vegetation patch. *Top right:* Classification after reducing the data to 25 bands using the proposed modified LLE. K-means is used for clustering, with the spatially consistent vector angle as the distance between projected pixels. *Bottom:* Same as first row for a different portion of the image.

specified to be non-calibrated and are set to zero, a few additional bands had to be excluded, they have more than 98% pixels with zero value. The 172/178 band image was reduced using the proposed modified locally linear embedding to 5, 10, 25, 50, and 100 projected bands. For each of these projections we varied the number of neighbors, using 5, 10, 25, 50, and 100. After the dimensionality reduction of the data, we used the classical K-means algorithm to cluster the pixels (recall that for this data, ground truth is not available). The proposed spatially coherent vector angle was used as the distance metric in the K-means algorithm. No ground truth is available for this data, therefore the number of classes were estimated from the visual clues. Specifically, for Figure 4, top, the three visible materials are the green vegetation patch, the mountain top rocks, and the plains. For Figure 4, bottom, the three visible materials are the mountain top, clouds, and the plains. Figure 4 shows the clustering results for projections onto 25 bands.

V. Concluding Remarks

In this letter we have introduced spatial coherence into dimensionality reduction and vector angle distance, and showed their contributions for the classification and clustering of hyper-spectral data. We effectively reduced the amount of data by more than 75%, while at the same time improving classification by more than 15%.

A lot of work remains to be done in the area of hyper-spectral classification and its connections with dimensionality reduction. First, in order to deal with large images, the use of classical works on out-of-sample dimensionality reduction need to be investigated. This will permit to work with a subset of the data while extending the learned dimensionality reduction map to the entire image. In the same flavor, it is also interesting to perform spatially-coherent LLE as here proposed on small segments of the hyper-spectral data, and then stitch them together in order to obtain the global map of the entire image. This idea of charting is in line with concepts presented in [2]. Results in this direction will be reported elsewhere.

Acknowledgments

This work was partially supported by NGA, ONR, DARPA, NSF, and the McKnight Foundation. The EO-1 satellite data was provided by the National Geospatial Intelligence Agency. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the National Geospatial-Intelligence Agency or the U.S. Government.

References

[1] http://aviris.jpl.nasa.gov/html/aviris.instrument.html

[2] C. M. Bachmann, T. L. Ainsworth, and R. A Fusina, "Exploiting manifold geometry in hyperspectral imagery," *IEEE Transactions Geoscience Remote Sensing* **43:3**, pp. 441-454, March 2005.

[3] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Proceedings of Advances in Neural Information Processing Systems* **14**, 2002.

[4] P. Bierwirth, R. Blewett, and D. Huston, "Finding new mineral prospects with HYMAP: Early results from a hyperspectral remote-sensing case study in west Pilbara," *ASGO Research Newsletter* **31**, pp. 29-31, November 1999-2000.

[5] A. Buades, B. Coll, and J. M. Morel, "On image denoising methods," *CMLA Preprint* (http://www.cmla.ens-cachan.fr/Cmla/), to appear in *SIAM Multiscale Modeling and Simulation*, 2006.

[6] C. I. Chang and Q. Du, "Interference and noise-adjusted principal components analysis," *IEEE Transactions Geoscience and Remote Sensing* **37**, pp. 2387, Sept. 1999.

[7] R. de Lange, M. van Tiland, and S. Dury, "The use of hyperspectral data in coastal zone vegetation monitoring," *EARSEL Proceedings* **3**, pp. 143-154, Feb. 2004.

[8] D. Donoho and C. Grimes, "When does ISOMAP recover the natural parameterization of families of articulated images?, http://www-stat.stanford.edu/∼donoho/reports.html, 2002.

[9] H. Du, H. Qi, X. Wang, R. Ramanath, and W. E. Snyder , "Band selection using independent component analysis for hyperspectral image processing," *Proc. 32nd Applied Imagery Pattern Recognition Workshop – AIPR 2003*, pp. 93-98, Washington DC, Oct. 2003.

[10] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," *IEEE International Conference on Computer Vision*, pp. 1033-1038, Corfu, Greece, Sept. 1999.

[11] J. C. Harsanyi and C. I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Transactions Geoscience Remote Sensing* **32:4**, pp. 779-785, 1994.

[12] D. C. Heinz and C. I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Transactions Geoscience Remote Sensing* **39:3**, pp. 529-544, March 2001.

[13] D. H. Kim and L. H. Finkel, "Hyperspectral image processing using locally linear embedding," *Proceedings of Neural Engineering-2003*, pp. 316-319, March 20-22 ,2003.

[14] H. Kwon and N. M. Nasrabadi, "Kernel orthogonal subspace projection for hyperspectral signal classification," *IEEE Transactions Geoscience Remote Sensing* **43:12**, pp. 2952-2962, Dec. 2005.

[15] M. Mahmoudi and G. Sapiro, "Fast image and video denoising via non-local means of similar neighborhoods," *IEEE Signal Processing Letters* **2**, pp. 839-842, December 2005.

[16] C. Rodarmel and J. Shan, "Principal Component analysis for hyperspectral image classification," *Surveying and Land Information Science*, **62:2**,pp. 115-122 June 2002.

[17] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science* **290**, pp. 2323-2326, 2000.

[18] K. Staenz, J. C. Deguise, J. Chen, H. McNaim, T. Szeredi, and M. McGovem, "The use of hyperspectral data for precision farming," *ISPRS Comission VI1 Symposium*, Budapest, Hungary, September 1-4, 1998.

[19] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science* **290**, pp. 2319-2323, 2000.