OBJECT DETECTION FROM HS/MS AND MULTI-PLATFORM REMOTE-SENSING IMAGERY BY THE INTEGRATION OF BIOLOGICALLY AND GEOMETRICALLY INSPIRED APPROACHES

Bo Wu[1], Yuan Zhou[1], Lin Yan[1], Jiangye Yuan[2], Ron Li[1], and DeLiang Wang[2]
[1]Mapping and GIS Laboratory, Department of Civil and Environmental Engineering and Geodetic Science
[2]Department of Computer Science and Engineering and Center for Cognitive Science
The Ohio State University, Columbus, OH 43210-1275
wu.573@osu.edu

ABSTRACT

This paper presents a system that integrates biologically and geometrically inspired approaches to detecting objects from hyperspectral and/or multispectral (HS/MS), multiscale, multiplatform imagery. First, dimensionality reduction methods are studied and used for hyperspectral dimensionality reduction. Then, a biologically inspired method, S-LEGION (Spatial - Locally Excitatory Globally Inhibitory Oscillator Network), is developed to perform object detection on the multispectral and dimension-reduced hyperspectral data, which provides rough object shapes. Thereafter, a geometrically inspired method, GAC (geometric active contour), is employed for refining object boundary detection on the high-resolution imagery based upon the initial object shapes provided by S-LEGION. A geospatial database is compiled and used for experimental analysis that includes data from a selected test site at Silver Lake in the Mojave Desert, California. Multispectral (Landsat TM 4-5) and hyperspectral (EO-1) satellite imagery, high-resolution satellite imagery (IKONOS), and descent images and ground stereo images are included in this database. This paper presents the first year results of a two-year research project.

INTRODUCTION

Over the last decades, there has been a remarkable increase in the number of remote-sensing sensors onboard various satellite-, aircraft-, and land vehicle-based platforms. Large volumes of panchromatic, multispectral, and hyperspectral data have been collected periodically. Fusion of these multi-platform remote-sensing data along with in situ observations from multiple sensors can help us to derive more information than is possible from a single sensor alone. Examples include detection of roads and buildings, determination of the composition of ground vegetation, and localization of mineral resources, as well as other application areas. However, the most detailed information, such as shape and spectral attributes, often cannot be derived precisely. Recent advances in biologically inspired methods involve segmenting patterns, materials, and objects, among other capabilities. Terman and Wang (1995) proposed locally excitatory globally inhibitory oscillator networks (LEGION) as a computational framework for image segmentation and object recognition. It has been shown analytically that LEGION networks can rapidly achieve both synchronization in a locally coupled oscillator group and desynchronization among different oscillator groups. LEGION has been successfully applied to segmenting grayscale images, medical images, and aerial images (Wang and Terman, 1997; Wang, 2005). On the other hand, geometrically inspired methods, such as level set theory and GAC (geometric active contour) models, have also been widely used in image segmentation and object detection (Osher and Sethian, 1988; Caselles et al., 1997; Niu, 2006). If the biologically inspired object detection methods can be combined with advanced geometry-based object detection techniques, a variety of object detection and recognition tasks in civilian, military, and intelligent applications can be significantly improved and speeded up.

This paper presents a system that integrates biologically and geometrically inspired approaches to detect objects from hyperspectral and/or multispectral (HS/MS), multiscale, multiplatform imagery. Dimensionality reduction methods are studied and used for hyperspectral dimensionality reduction. A biologically inspired method, S-LEGION (Spatial - LEGION), is developed to perform object detection on the multispectral and dimension-reduced hyperspectral data, which provides rough object shapes. Then, a geometrically inspired method, GAC (geometric active contour), is employed for refining object boundary detection on the high-resolution imagery based on the initial object shapes provided by S-LEGION. This research is funded by a NGA University Research Initiatives project. This paper presents the results of the first year of the project, mainly summarizing the architecture of the integrated system for object detection and hyperspectral dimensionality reduction.

AN INTEGRATED SYSTEM FOR OBJECT DETECTION

The architecture and concept of the system integrating biologically and geometrically inspired object detection methods are illustrated in Figure 1. First, satellite and airborne HS/MS data will be processed using an S-LEGION algorithm for image segmentation in order to study regional and contextual information about the entire site. Dimensionality reduction methods will be employed for hyperspectral dimensionality reduction. Spectral compositions of HS/MS images will help to quickly identify candidate regions for detecting objects of interest. Supported by high-resolution satellite images, multiscale descent images, and ground images, a GAC model will then be applied to each small region of interest detected by S-LEGION for improved boundary extraction and shape reconstruction. The extracted information will serve as input to a final object recognition method using shape-based and spectral-based techniques. During the entire process, a multiplatform sensor modeler will be applied to support precision mulitsensor and multiplatform image analysis.

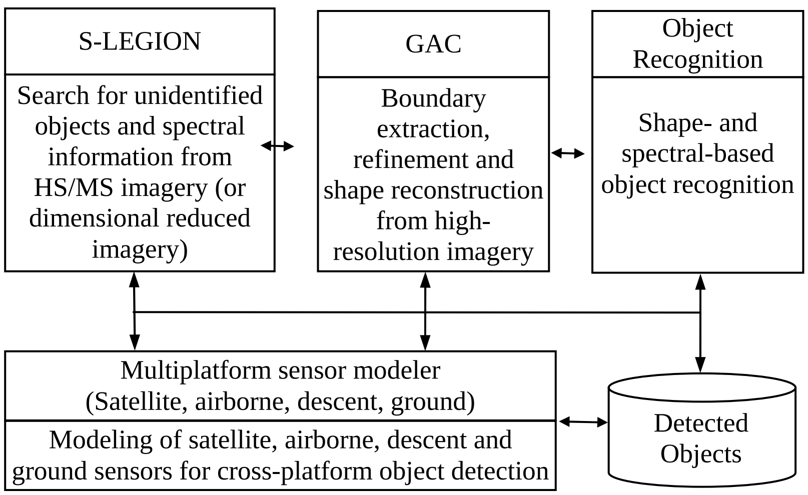


Figure 1. Integrated system for biologically/geometrically-inspired object detection.

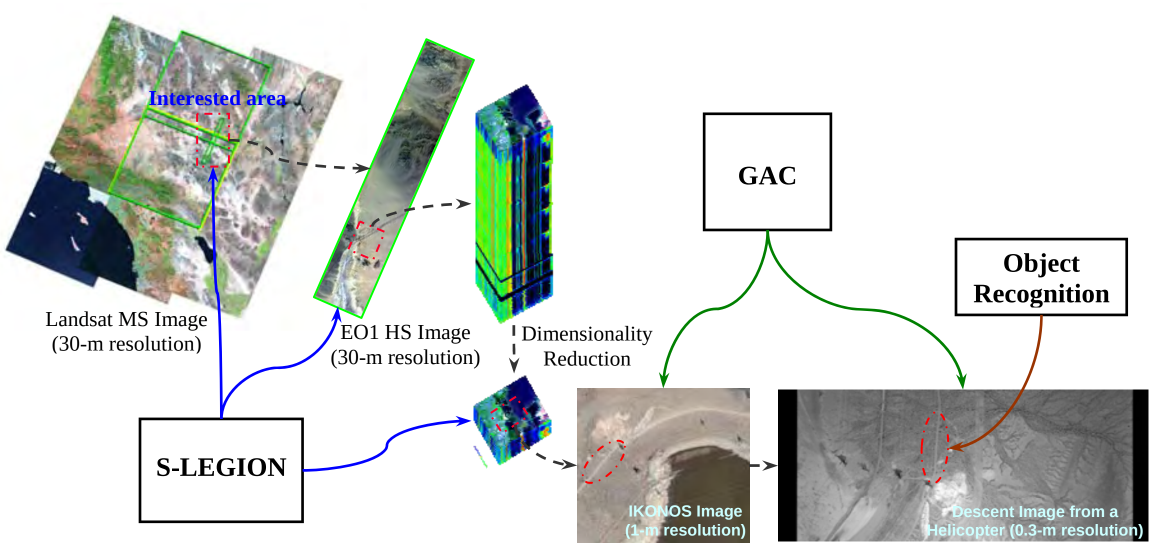


Figure 2. Conceptualization of the integrated system for object detection using the multi-platform data collected at Silver Lake in the Mojave Desert, CA.

In this investigation, we will apply the proposed approach to data compiled into a geospatial database using the data collected from field tests at Silver Lake in the Mojave Desert. The geospatial database includes multi-scale images collected from satellite, airborne, descent (helicopter), and in situ (land vehicle and field robot) images as well as GPS control points. It has been collected and maintained since 1998. The descent images represent a sequence of multiscale images that can be used to test the capability of scale invariant object recognition. The Silver Lake test site consists of desert terrain with scattered bushes, a dry lake, highways, paved and unpaved roads, utility lines, a helicopter port, and small housing complexes. Figure 2 illustrates the concept of the integrated system for road detection using the Silver Lake data. This geospatial database will contribute significantly to system development and validation.

HYPERSPECTRAL DIMENSIONALITY REDUCTION

Object detection methods (such as LEGION and GAC mentioned above) have been successfully applied to process images including medical images, close-range images, and airborne/satellite images. However, problems will arise from transplanting these algorithms to HS/MS remote-sensing imagery. Compared with the single-band image, the new data form (especially the hyperspectral image) has a much higher number of dimensions in spectral space. This increase in dimensions results in a rapid increase in computational costs and a reduction in classifier performance. This is called the “Hughes Phenomenon” (Hughes, 1968). Moreover, there is information redundancy in the high-dimensional space, such as the high correlation between those spectral bands close to each other and the noise from atmospheric absorption. This information redundancy wastes computation time and depresses accuracy in the image processing. Therefore, dimensionality reduction for hyperspectral imagery is necessary before it can be used for the subsequent object detection.

Dimensionality Reduction Methods

Dimensionality reduction maps a high-dimensional space onto a space with fewer dimensions, while the data in the original space can still be fully represented. It reduces the impact of the “Hughes Phenomenon” and also reduces redundant information, thus raising the efficiency of the data processing.

Many dimensionality reduction methods have been presented in the past; Principal Component Analysis (Hotelling, 1933), Linear Discriminant Analysis (Fisher, 1936), and Maximum Noise Fraction (Green et al., 1988) have been efficiently applied in compressing the HS/MS images. These methods share the common employment of eigenvalue-based linear approximation to retrieve the original spectral space. However, from a spectral point of view, spectra characteristics are treated as linear features in these approaches, while their nonlinear features are ignored. However, in certain circumstances, the nonlinearities of hyperspectral data can be the major properties in spectral space (Bachmann et al., 2004). Li et al. (2005) proved that the nonlinear spectral inverse model is more accurate than the linear method.

In recent years, Manifold Learning techniques have been introduced to model the nonlinear features (manifold) of high-dimensional data and to project the manifold onto low-dimensional space whereby the nonlinear properties of the data could be well preserved during the data compression. Basically, these techniques could be divided into two groups: the first one preserving the global properties of the data and the second one focusing on the local properties of the data (van der Maaten et al., 2007). The representative methods for the first group include ISOmaps (Tenenbaum et al., 2000), Kernel PCA (Scholkopf et al., 1998), and Diffusion Maps (Lafon and Lee, 2006). The second group includes Local Linear Embedding (Roweis, 2000; Han and Goodenough, 2005), Laplacian Eigenmap (Belkin, 2003), and Local Tangent Space Alignment (Zhang et al., 2002). Compared with the global properties preserved in Manifold Learning methods, local nonlinear techniques for dimensionality reduction are based on solely preserving the properties of small neighborhoods around the data points of interest. This satisfies the principles of object detection within a local neighborhood. This research mainly examines the following three local nonlinear techniques.

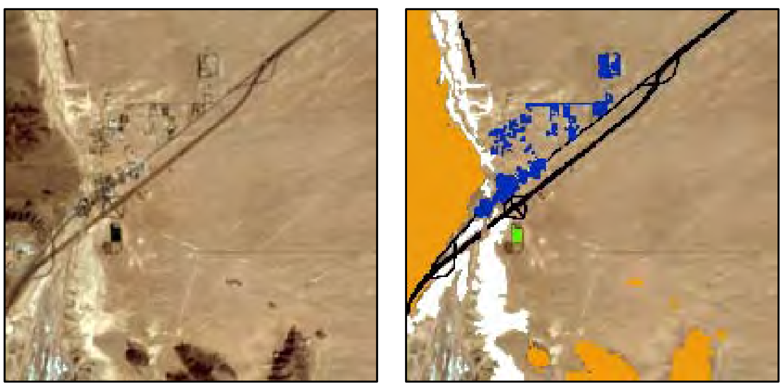
Local Linear Embedding (LLE) is the first local Manifold Learning technique introduced for dimensionality reduction (Roweis, 2000; Han and Goodenough, 2005). In LLE, the points neighboring the data points of interest are found and saved in a neighborhood graph. The local geometry of the data points is estimated by the reconstruction weights using a cost function. In this function, the reconstruction weights are subject to the constraint that the difference between the linear combination of the weighted neighbor points and the center point should be minimized. The data points will be projected into a new low-dimensional space based on minimizing the linear combination of the point and its neighbor points in the new space, weighted by the reconstruction weights.

Laplacian Eigenmap (LE) calculates the weight between data points and their neighborhood using the nearest neighborhood methods (Belkin, 2003). The nearest neighbor of the data point will contribute the most. The weights are used as edges connecting each point with its neighbors in a graph. By using spectral graph theory, the projection from a high-dimensional space to a low-dimensional space is defined as an Eigen problem. A Laplacian matrix is derived from the connected weight edges of the graph. The data points in the low-dimensional space are generated by the linear combination of the largest eigenvectors of the Laplacian matrix. The number of eigenvector is the same as the number of the dimensions in the low-dimensional space.

Local Tangent Space Alignment (LTSA) uses the local tangent space of each data point to describe the local properties of the high-dimensional space (Zhang et al., 2002). In LTSA, it is assumed that if there is local linearity in the manifold surface, the data points in the high-dimensional space and the corresponding low-dimensional space could be mapped to the same local tangent space. In other words, LTSA simultaneously searches for the local tangent space of the high-dimensional data and the low-dimensional data representations, and the mapping relationships between the low-dimensional data points and the local tangent space of the high-dimensional data.

Experimental Results

This paper compared the above-mentioned three Manifold Learning methods (LLE, LE, and LTSA) using hyperspectral data (Figure 3) collected at Silver Lake. This hyperspectral data was selected from a 30-m resolution EO1 Hyperion Hyperspectral Level 1GST product that was acquired in October 2003. The image size for the selected experimental area is 200 x 200 pixels (Figure 3a). It has 242 bands (from 447.17 nm to 2577.08 nm) in spectral space. Pre-processing of spectral and geometric correction was performed using the methods proposed by Beck (2003) and Datt et al. (2003). After removal of atmospheric absorption effects, 156 bands were left for later dimensionality reduction experiments. Using the Maximum Likelihood Estimation method (Levina and Bickel, 2004), the intrinsic dimensions of the input data were estimated as 15. In our experiments, 40,000 data points (200 x 200) with 156 dimensions are used as input data for the dimensionality reduction process using the LLE, LE, and LTSA methods; the output is the reduced data with 15 dimensions. The time-consumption ratio for these three methods is 15(LLE):2(LE):10(LTSA), showing that the LE method is the most efficient.

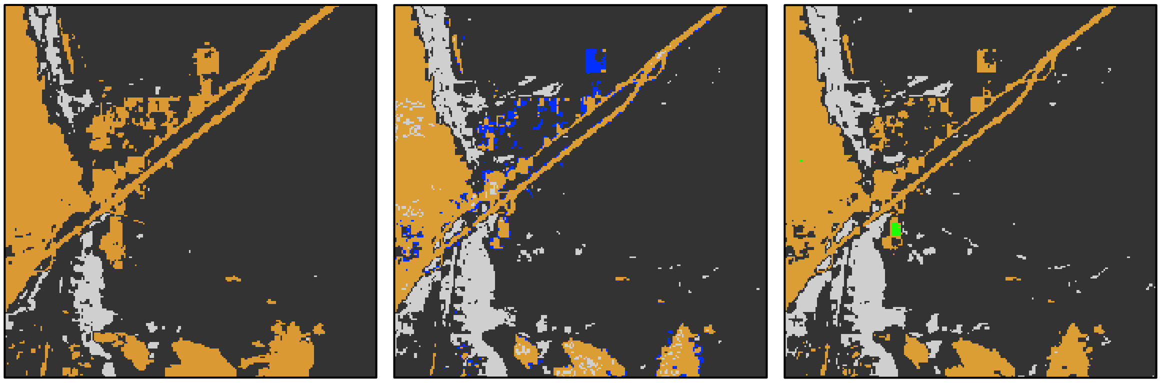


a) b)

Figure 3. Experimental data for dimensionality reduction: a) a pseudocolor EO1 hyperspectral image of 200 X 200 pixels, and b) manually digitized objects as ground truth (brown denotes hills, white lakebed, black asphalt roads, blue buildings, and green vegetation).

To evaluate the results of these three dimensionality reduction methods, five categories of objects were manually digitized and labeled on the pseudocolor image. These objects include hills, lakebed, asphalt road, buildings and vegetation as illustrated in Figure 3b. The coverages for these objects are 4781 pixels for the hills (brown in Figure 3b), 2353 pixels for the lakebed (white), 950 pixels for asphalt roads (black), 964 pixels for buildings (blue), and 39 pixels for vegetation (green). These objects were used as ground truth in our experiment.

After dimensionality reduction using the LLE, LE, and LTSA methods, 15 dimensions data were obtained for each of these methods. Then an unsupervised classifier, the K-nearest neighborhood (KNN) method, was used to perform classification for the dimension-reduced data. Each dataset was separated into eight clusters. Those clusters corresponding to the manually digitized ground-truth objects were assigned their same color; the remaining clusters were all colored black and treated as background. These classification results are illustrated in Figure 4.



a) b) c)

Figure 4. Classification results of the dimension-reduced data from the a) LLE, b) LE, and c) LTSA methods.

As shown in Figure 4, we found that the lakebed could be separated in the classification results from all three dimensionality reduction methods. However, the asphalt roads and hills are mixed up in all three cases. In a field spectral survey performed at Silver Lake in October 2008, we found that these two types of objects are inseparable from the spectral point of view due to their high correlation of spectrum features (this could be the consequence of using the materials collected from the hills to construct the roads). Therefore, we merged the hills and asphalt roads in the following analysis of the classification results.

To analyze the classification results, we compared the numbers of pixels classified in each of the clusters based on the three different dimensionality reduction methods with the numbers of pixels from the corresponding manually digitized ground truth. The results are listed in Table 1. It can be seen that the LTSA provides the best average classification rate (80.02%), while the LE has the lowest average classification rate (73.47%). Buildings can be isolated from other objects in the results based on LE, while vegetation can be isolated by the classifier based on LTSA. In general, it is believed that LTSA performs better than the other two dimensionality reduction methods, LE and LLE.

Table 1. Classification results based on the LLE, LE, and LTSA methods.

Methods / Objects	LLE			LE			LTSA		
	Pixels from ground truth	Classified pixels	Rate	Pixels from ground truth	Classified pixels	Rate	Pixels from ground truth	Classified pixels	Rate
Hills and Road	5465	7188	76.03%	4946	5907	83.73%	5127	5467	93.78%
Lakebed	1296	1479	87.63%	2053	3528	58.19%	2152	3633	59.23%
Buildings	N/A	N/A	N/A	349	567	61.55%	N/A	N/A	N/A
Vegetation	N/A	N/A	N/A	N/A	N/A	N/A	30	34	88.24%
Average Rate	78.01%			73.47%			80.02%		

Neighborhood Distortion Index for Performance Evaluation of Dimensionality Reduction

The above experimental analysis mainly uses statistics to evaluate the performance of the different dimensionality reduction methods. However, these statistics are largely dependent on the performance of the classifier adopted in the processing. Consequently, this statistic analysis may not be capable of fully studying the capabilities of these various methods. This paper proposes a criterion, the Neighborhood Distortion Index (NDI), for evaluating the performance of the dimensionality reduction methods.

We assume that the ideal dimensionality reduction method should fully preserve the topological relationships between the data points during the reduction in spectral space. This means that the same clustering results could be obtained from either the original high-dimensional dataset or the dimension-reduced dataset. If there are some inconsistencies, then we consider them to be distortions due to the dimensionality reduction. The NDI has been developed to evaluate the type of distortion caused by the dimensionality reduction methods.

For any pixel a in a hyperspectral image data set, there will be a vector V_a in the spectral space. The n neighboring pixels around a will also have vectors V_{bi} (i = 1,2,...,n) in the spectral space (Figure 5). The NDI D_a for a can be calculated using the following equation:

$$D_a = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\alpha_i - \alpha_i'}{\alpha_i}\right) \qquad (1)$$

where α_i is the intersection angle between vector V_a and V_{bi} in the original spectral space and α_i' is the corresponding angle in the reduced low-dimension space. We use the intersection angles of the vectors to describe their topological relationship. After the NDIs for all the pixels in the hyperspectral image have been calculated, an NDI map can be generated to illustrate the overall distortion of the results from the dimensionality reduction.

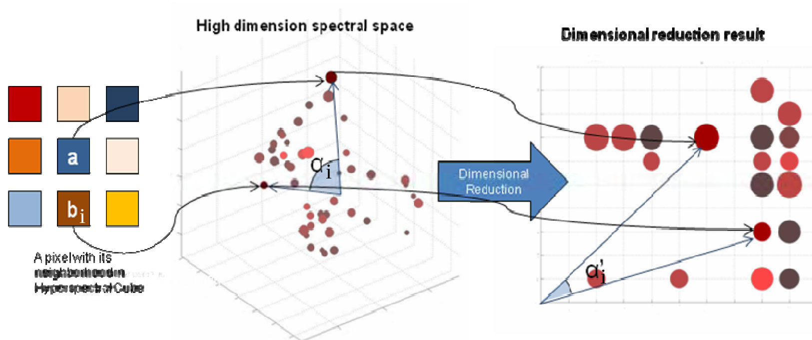
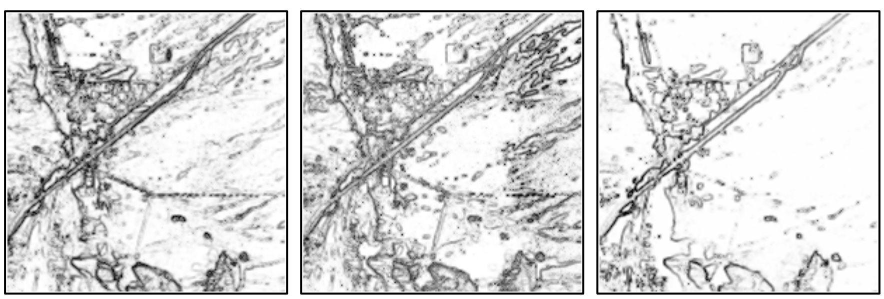


Figure 5. The topological relationships between the high-dimensional data and the dimension-reduced data.

Figure 6 shows the NDI maps of the dimensionality reduction results from LLE, LE and LTSA, respectively. In these maps, the pixels with larger NDIs have darker gray values. The average NDI for the LLE, LE, and LTSA NDI maps are 0.91, 1.43 and 0.32, respectively. This means that LTSA performs best for preserving topological relationships for dimensionality reduction. These results are consistent with those results obtained from the previous analysis by comparing with the ground truth.



a) b) c)

Figure 6. NDI maps of the results of dimensionality reduction from: a) LLE, b) LE, and c) LTSA.

DISCUSSION AND CONCLUSIONS

This paper investigates a system that integrates biologically and geometrically inspired approaches to detect objects from HS/MS, multiscale and multiplatform images. The architecture and concept of the integrated system have been studied and illustrated.

As the first step of the proposed system, this paper studies the dimensionality reduction methods for hyperspectral dimensionality reduction. Using the hyperspectral data collected at Silver Lake, three dimensionality reduction methods including LLE, LE, and LTSA, were used for dimensionality reduction. To compare the performance of these methods, a Neighborhood Distortion Index is developed. By experimental statistic comparison and analysis using NDI maps, LTSA has the best capability to preserve the features in high-dimensional dataset reduction.

This paper summarizes the preliminary results of our research. Our future works will focus on the further development of the S-LEGION and GAC to fully implement the proposed integrated system for biologically/geometrically-inspired object detection.

ACKNOWLEDGEMENT

This work was supported by an NGA University Research Initiatives grant (HM 1582-07-1-2027).

REFERENCES

Bachmann, C.M., T.L. Ainsworth, and R.A. Fusina, 2004. Modeling data manifold geometry in hyperspectral imagery, *Proceedings of 2004 IEEE Geoscience and Remote Sensing Symposium*, *5*: 3203-3206.

Beck, R., 2003. EO-1 user guide v.2.3, Available at http://eo1.usgs.gov/documents/.

Belkin, M., and P. Niyogi, 2003. Laplacian eigenmaps for dimensionality reduction and data representation, *Neural Computation*, *15*(6):1373-1396.

Caselles, V., R. Kimmel, and G. Sapiro, 1997. Geodesic active contours, *International Journal of Computer Vision*, *22*(1):61-79.

Datt, B., T.R. McVicar, T.G. van Niel, D.L.B. Jupp, J.S.Pearlman, 2003. Preprocessing EO-1 Hyperion hyperspectral data to support the application of agricultural indexes, *IEEE Transactions on Geoscience and Remote Sensing*, *41*(6, Part 1):1246-1259.

Fisher, R.A., 1936. The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, *7*:179-188.

Green, A.A., M. Berman, P. Switzer, and M.D. Craig, 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal, *IEEE Transactions on Geoscience and Remote Sensing*, *26*:65-74.

Han, T., and D.G. Goodenough, 2005. Nonlinear feature extraction of hyperspectral data based on locally linear embedding (LLE), *Proceedings of 2005 Geoscience and Remote Sensing Symposium*, *2*:1237-1240.

Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology*, *24*:417-441.

Hughes, G.F., 1968. On the mean accuracy of statistical pattern recognizers, *IEEE Transactions on Information Theory*, *IT-14*:55-63.

Lafon, S., and A.B. Lee, 2006. Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(9):1393-1403.

Levina, E., and P.J. Bickel, 2004. Maximum Likelihood Estimation of Intrinsic Dimension, Neural Information Processing Systems: NIPS, Vancouver, CA, December, 2004.

Li, D., S. Fang, and Y. Zhou, 2005. A technique for inversion of water quality parameters in the East Lake based on spectral reflectance measurement, *Proceedings of SPIE, the International Society for Optical Engineering*, *6043*(2):60430M.1-60430M.7.

Niu, X., 2006. A semi-automatic framework for highway extraction and vehicle detection based on a geometric deformable model, *ISPRS Journal of Photogrammetry and Remote Sensing*, *61*(3-4):170-186.

Osher, S., and J. Sethian, 1988. Fronts propagating with curvature-dependent speed: algorithms based on the Hamilton-Jacobi Formulation, *Journal of Computer Physics*, *79*(1):12-49.

Roweis, S.T., 2000. Nonlinear dimensionality reduction by locally linear embedding, *Science*, *290*(5500):2323-2326.

Scholkopf, B., A.J. Smola, and K.-R. Muller, 1998. Nonlinear component analysis as a kernel eigenvalue problem, *Neural Computation*, *10*(5):1299-1319.

Tenenbaum, J.B., V. de Silva, and J.C. Langford, 2000. A global geometric framework for nonlinear dimensionality reduction, *Science*, *290*(5500):2319–2323.

Terman, D., and D.L. Wang, 1995. Global competition and local cooperation in a network of neural oscillators, *Physica D*. *81*:148-176.

van der Maaten, L.J.P., E.O. Postma, and H.J. van den Herik, 2007. Dimensionality reduction: A comparative review, Available at http://www.cs.unimaas.nl/l.vandermaaten/dr/DR_draft.pdf.

Wang, D.L., and D. Terman, 1997. Image segmentation based on oscillatory correlation, *Neural Computation*, *9*: 805-836.

Wang, D.L., 2005. The time dimension for scene analysis, *IEEE Transactions on Neural Networks*, *16*: 1401-1426.

Zhang, Z., and H. Zha, 2004. Principal manifolds and nonlinear dimensionality reduction via local tangent space alignment, *SIAM journal of Scientific Computing*, *26*(1): 313-338.