TREE DISEASE DETECTION BASED ON HYPERSPECTRAL MANIFOLD LEARNING OF PHENOLOGICAL TRANSITION ON FOREST

Kuniaki Uto[*1] and Yukio Kosugi [2]
[1]Assistant Professor Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, G2-16, 4259, Nagatsuta-cho, Midori-ku, Yokohama 2268502, Japan; Tel: +81-45-9245484
Email: uto@pms.titech.ac.jp

[2]Professor Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, G2-16, 4259, Nagatsuta-cho, Midori-ku, Yokohama 2268502, Japan; Tel: +81-45-9245466
Email: kosugi@pms.titech.ac.jp

KEY WORDS: Hyperspectral data, Manifold, Tree disease, Forest, Phenology

ABSTRACT: It is a challenging task to extract tree disease from a wide spectral variety of autumnal tints, in spite the fact that the prevention and extermination based on early detection of affected tree is the only measure to keep infectious tree disease out at the borders. It is expected that the hyperspectral subspace of healthy trees forms a smooth manifold spanned by features, e.g. pigments, leaf structure and water, because the phenological transition in hyperspectral space is reflected by continuous change in the features. In this study, we investigate hyperspectral subspace structure of phonological distribution of healthy forest in order to detect tree disease, e.g. Japanese oak wilt. The aim of this study is the detection of local coordinates by which the anomaliness of affected trees is remarkable. The processing flow is composed of (1) Definition of local tangent spaces, (2) Anomaly data projection onto the subspace defined by the local tangent spaces, (3) Linear subset extraction based on clustering of the projected data, and (4) Local coordinate extraction of the clustered subsets. Fixed point observations of hyperspectral images of healthy deciduous forests from mid-summer to winter were carried out by a VNIR hyperspectral imaging sensor. VNIR hyperspectral images of oak wilt trees as affected targets were collected in mid-summer as well. Experimental results show that the proposed method extracts local linear coordinates to enhance the anomaly features.

1. INTRODUCTION

Infectious wilt diseases, e.g. oak wilt (Kamata, 2002) and pine wilt, are becoming critical threats worldwide. A demand for automatic mapping method based on remotely sensed forest image is growing because the prevention and extermination based on early detection of affected tree is the only measure to keep infectious tree disease out at the borders. Since the optical properties of forest reflect the change in pigments and structure, hyperspectral sensing is a promising tool for the disease detection (Uto et al., 2008).

Due to the fact that the phonological transition in spectral features is caused by continuous changes in quantities of pigments, e.g. chlorophyll, carotenoid and water, and structure in leaf and canopy scale, it is expected that the temporal-spectral distribution of normal forest is modeled by a low dimensional nonlinear manifold, i.e. it is difficult to apply conventional linear discriminant analyses in which the discriminant boundaries are optimized based on the assumption of equal class covariances. Consequently, it is indispensable to investigate the spectral subspace structure of normal forest in order to achieve precise disease tree detection from phonological spectral variation as background distribution.

Manifold learning provides a way of understanding the phonological transition of forest in hyperspectral space. In a variety of methods proposed for manifold learning (Roweis and Saul, 2000; Tenenbaum et al., 2000; Zhang and Zha, 2002), smooth coordinates on the manifold are computed with constraints of global uniformity of the dimensionality and local linearity of the topological structure. By local tangent space alignment (Zhang and Zha, 2002), the manifold structure is optimized based on the local tangent space alignment by which global intrinsic coordinates of the manifold are derived while the local structures in small neighborhoods are preserved. In those manifold learning methods, the intrinsic structure of the manifold is extracted without referring heterogeneous data, e.g. wilt trees, outside of the manifold. Therefore, manifold learning referring to outside singular data is highly required to achieve precise anomaly detection.

In this paper, a novel manifold learning method based on the relationship between the hyperspectral manifold and the outside anomaly data is proposed to estimate local linear coordinates characterizing the anomaly features of the affected data.

2. METHODOLOGY

2.1 Overview of Algorithm

The aim of the proposed method is the extraction of linear subsets and local coordinates characterizing the anomaly features of the affected data. The processing flow is composed of (1) Definition of local tangent spaces, (2) Anomaly data projection onto the subspace defined by the local tangent spaces, (3) Linear subset extraction based on clustering of the projected data, and (4) Local coordinate extraction of the clustered subsets.

2.2 Definition of Local Tangent Space

Theoretically, the local tangent space is defined by the Jacobi matrix of the unknown manifold function. The local coordinate, matrix Q_i, i=1,…N of d columns, is approximated by d left singular vectors of $X_i(I\text{-}ee^T/k)$, i=1,..,N, corresponding to the d largest singular values (Zhang and Zha, 2002), where d is a given dimensionality, X_i is a matrix consisting of the k-nearest neighbors of a spectral point $\boldsymbol{x}_i$, I is an identity matrix and e is an N-dimensional column vector of all ones. N is the number of data.

2.3 Projection onto the Subspace Defined by the Local Tangent Space

The spectral change from normal to affected trees is a state transition between normal state and affected state, so that it is expected that the transition in spectral space is discontinuous and the affected trees are placed outside of the manifold. However, in order to discover the structure of the manifold for tree disease detection, it is encouraging to make good use of the relationship between the normal local subspace and the anomaly data based on the local coordinates. The coordinate, $\boldsymbol{a}_i^P$, of the affected data is calculated by the orthogonal projection onto the local tangent space based on projection matrix, P_i, defined by the local coordinates in Equation (1), where P_i is defined by P_i=$Q_iQ_i^T$, $\overline{\boldsymbol{a}}$ is a mean of anomaly data $\boldsymbol{a}$. N is the number of data

$$\boldsymbol{a}_i^P = P_i(\overline{\boldsymbol{a}} - X_i e/k) + X_i e/k\,, i = 1, \dots, N \tag{1}$$

2.4 Linear Subset Extraction

The pigments and structure transition of normal leaves is not simultaneous and linear, e.g. the decomposition of chlorophyll at early stage of tint is followed by the decomposition of carotenoid at late stage of tint. Therefore, it is expected that globally continuous manifold consists of subsets of the subspace in which different parameters are dominant coordinates of the subspaces. By assuming the consistent and low curvature within each subset, i.e. linearity of the subspace, the projected data belonging to same subset are spatially clustered. The subsets of the clusters are classified by Gaussian mixture model (GMM) based on EM algorithm (McLachlan and Peel, 2000). GMM is a probabilistic model defined in Equation (2).

$$p(\boldsymbol{x}) = \sum_{j=1}^{M} w_j\, g\left(\boldsymbol{x}\middle|\mu_j, \Sigma_j\right), \tag{2}$$

where $g\left(x\middle|\mu_j, \Sigma_j\right)$ is a Gaussian distribution function of $\boldsymbol{x}$ with a mean μ_j and a variance-covariance matrix Σ_j. w_j is a weight of $g\left(\boldsymbol{x}\middle|\mu_j, \Sigma_j\right)$, and M is the number of mixture components. w_j, μ_j, and Σ_j are optimized by EM algorithm based on the set of the projected data $\boldsymbol{a}_i^P, i = 1 \dots N$. Clusters with high mixture weights correspond to the spatially clustered subsets. The clusters in manifold, C_j, and projected data, C^p_j, are defined in Equation (3) and (4).

$$C_j = \{\boldsymbol{x}_i | j = \arg_k \max g(\boldsymbol{x}_i|\mu_k, \Sigma_k)\} \tag{3}$$

$$C_j^P = \{\boldsymbol{a}_i^P | j = \arg_k \max g(\boldsymbol{a}_i^P|\mu_k, \Sigma_k)\} \tag{4}$$

2.5 Local Coordinate Extraction

Since it is expected that the manifold are decomposed to subsets with low curvatures, the local coordinates of the subsets are approximated by linear bases. The dominant local linear coordinates $\boldsymbol{v}_i$ for the anomaly detection are defined by Equation (5).

$$\boldsymbol{v}_i = \overline{C_J^P} - \overline{C_J} \tag{5}$$

where $\overline{C_j}$ and $\overline{C_j^P}$ are means of C_j and C_j^P respectively. For a given spectral datum $\boldsymbol{x}$, a score $s_i(\boldsymbol{x})$ based on the local linear coordinate vector $\boldsymbol{v}_i$ is given by $s_i(\boldsymbol{x}) = x^T \cdot \boldsymbol{v}_i$.

2.3 Simulation

Simulation results based on 1-D manifold sample points in 2-D space are shown in Figure 1. The green and red points in Figure 1(a) correspond to 1-D normal manifold and anomaly data respectively. 200 manifold data are generated by a quadratic polynomial, $f(\tau) = (\tau, -5(\tau-1)^2 + 3), \tau \in [0\ 2]$, with additive Gaussian noise of variance 0.2. The anomaly data are normal distribution of center (1.5, -1.0) and variance 0.2. Projected data on the subspace defined by local tangent space are plotted in Figure 1(b). The neighborhood size k is 20. The magenta, cyan and black points in Figure 1(b) are center positions of three components in Gaussian mixture model. Figure 1(d) shows clusters corresponding to the points with the first and second mixture weights in Figure 1(b). Green and red points in Figure 1(c) are points corresponding to the two classes shown in Figure 1(d).

The projection method is applied to 2-D manifold sample points in 3-D space (Figure 2). 20000 points ware generated by a cubic polynomial, $f(\tau_1, \tau_2) = (\tau_1, \tau_2, 10\tau_1^3 + 2\tau_1^2 - 10\tau_1), \tau_1, \tau_2 \in [-1.5\ 1.5]$, with additive Gaussian noise of variance 1.0. The singular data are normal distribution of center (-0.5, 0, -10.0) and variance 0.2. The blue and red points in Figure 2(a) correspond to 2-D normal manifold and anomaly data respectively. Green, red and blue points in Figure 2(b) are classified points.

In Figure 1(c) and 2(b), it is verified that subsets with low curvature in the manifold are clustered based on the proposed method, whereas the intermediate data with high curvature are excluded.

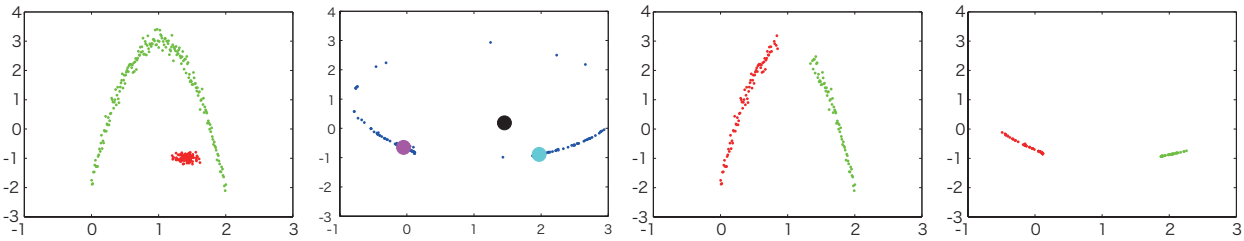
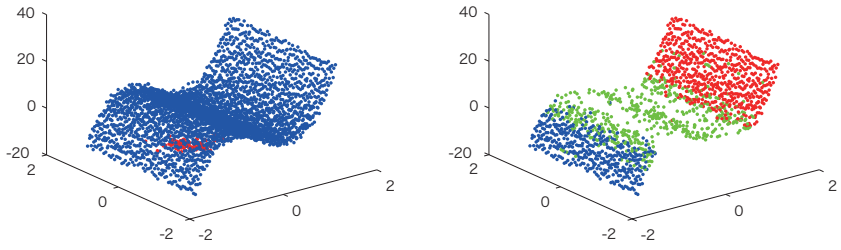


Figure 1. (a) Simulated 1-D manifold data and anomaly data distribution, (b) Projected data on the subspace defined by local tangent space and center positions of three components in Gaussian mixture model, (c) Data distribution of classified data based on the Gaussian mixture model in manifold space, and (d) Data distribution of classified data based on the Gaussian mixture model in projected space.



Fugire 2. (a) Simulated 2-D manifold data and anomaly data distribution, (b) Data distribution of classified data based on the Gaussian mixture model in manifold space

3. EXPERIMENTAL RESULTS

3.1 Hyperspectral Data

Hyperspectral data of healthy and affected trees were extracted from hyperspectral images of forests. The hyperspectral images were collected by a hyperspectral line array sensor. The spectral range is from visible to near-infrared, 500-900nm, with spectral resolution of 5nm and 81 bands. The spatial resolution is converted to 2.5

meters by spatial smoothing. The collected spectral data were normalized (Wu, 2004) to reduce the effect of brightness variation caused by the leaf density and illumination angle.

Normal spectral transition of deciduous forest were collected by fixed observations from Oct. 11 to Dec. 18 in 2007 in Yokohama, Japan. Spectral data of oak wilt trees were collected on Sep. 3 in 2007 and Aug. 27 in 2010 in Yamagata, Japan. The oak wilt trees were identified based on the existence of leaf wilt and frass on the trunk in the field surveys. The numbers of normal data and oak wilt data are 16,442 and 839, respectively.

3.2 Verification of Singularity of Affected Trees in Hyperspectral Space under the Framework of Anomaly Detection

Although there are singular characteristics in wilt leaves, e.g. withering and decay, it is visually difficult to distinguish affected trees from healthy autumnal tints, esp. late stage tints, in remote sensing scale. In this section, prior to the construction of hyperspectral manifold, the hyperspectral singularity of affected trees is evaluated based on anomaly detection.

On the assumption that the hyperspectral data of healthy trees are distributed within a low dimensional subspace, the singularity of affected trees is evaluated by the distance from the manifold. In this section, the singularity is verified under the framework of anomaly detection, instead of constructing local coordinates using manifold learning method. Originally, hyperspectral anomaly detection methods have been proposed to detect unknown anomaly object from normal background by evaluating the Mahalanobis distance between the pixel of interest and the background class in the hyperspectral space, i.e. Reed-Xiaoli (RX) algorithm (Reed and Yu, 1990), or in the feature space, i.e. kernel RX algorithm (Banerjee et al., 2006), and kernel-base support vector data description (K-SVDD) (Heesung and Nasrabadi, 2005).

Histograms of RX scores and kernel RX scores of healthy trees and oak wilt trees are shown in Figure 3. Gaussian RBF kernel with experimentally adjusted variance 5.0 is used for the kernel RX algorithm. It is verified that there is a significant gap between the manifold of healthy trees and oak wilt class. The definite difference between the healthy trees and oak wilt trees in Figure 3(a) indicates that the oak wilt class exists outside of convex hull of healthy tree classes.

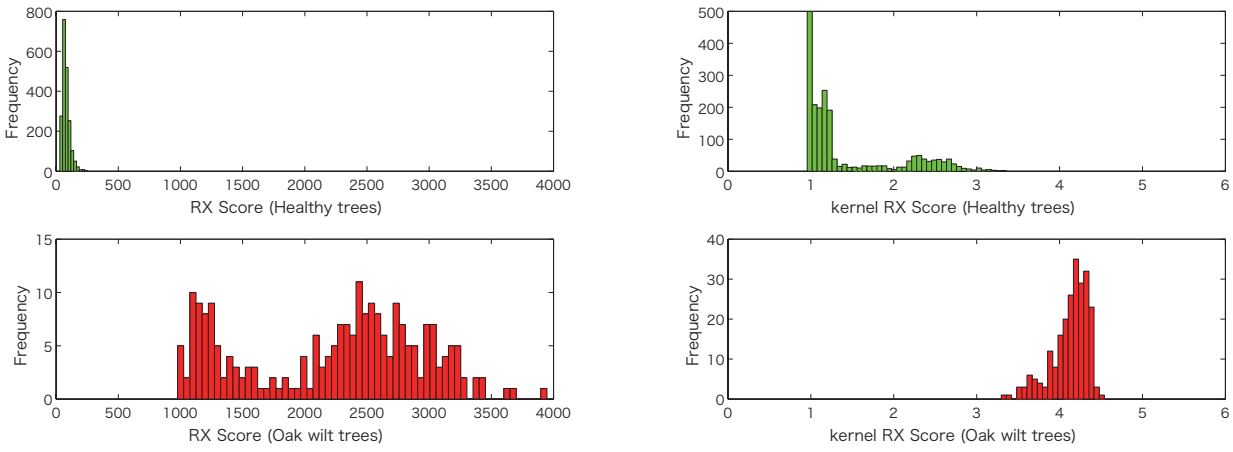


Figure 3. (a) Histograms of RX scores, top: Healthy trees, bottom: Oak wilt, (b) Histograms of kernel RX scores, top: Healthy trees, bottom: Oak wilt.

3.3 Manifold Learning of Normal Forest Phenology

Hyperapectral data of normal trees are decomposed to a set of low dimensional linear subspace by the proposed method based on hyperspectral data of anomaly oak wilt trees. In Figure 4(a), blue and red points are 2-D normal leaf data and oak wilt data on 530-786 [nm] normalized reflectance space respectively. Green and magenta points in Figure 4(b) are classified points corresponding to clusters with the first and second highest mixture weights. It is confirmed that the green and magenta clusters are correlated with early stage of tint and late stage of tint. Figure 4(c) and (d) show the 3-D distribution on 545-680-786[nm] normalized reflectance space. The neighborhood size k and the number of components of GMM are 20 and 5 respectively.

Figure 5(a) and (b) show histograms of scores based on local coordinate vectors of clusters 1 and 2. It is verified that anomaly oak wilt data are distributed significantly apart from cluster 1 and 2. In the bottom distributions in Figure 5(a) and (b), the score distributions of whole normal trees are widely distributed so that there is overlap between the distribution of whole normal trees and oak wilt, e.g. Figure 5(a) middle and bottom. The results indicate that data analysis based on a set of local coordinates is indispensable for precise detection. In Figure 6(a) and (b), histograms of scores based on the first principal vectors in cluster 1 and 2, instead of the local linear vector defined by Equation (4), are shown. There is distribution overlap between the top and middle histograms in Figure 6(b), which indicates that optimal local coordinate for anomaly detection is not equivalent to the principal vector. Red and black lines in Figure 4(b) show the local coordinate vectors (solid lines) and the first principal vectors (dotted lines).

4. CONCLUSION

A new manifold learning method for anomaly detection is introduced. The anomaly features of oak wilt spectral data are enhanced by the estimated local linear coordinates.

GMM is applied to clustering of projected data and the experimental results show satisfactory performance. However, the selection of the number of the mixture components is so critical that the adaptive clustering method will be required. In optical remote sensing field, normalized difference indices, e.g. NDVI (Rouse et al.) and NDWI (Gao, 1996), realize successful results in various applications, despite the simple definitions approximated by projections onto linear vectors as well. In order to deal with more complicated pattern recognition problems, the combination of the simple indices, as proposed in this study, could be one of the reliable approaches.

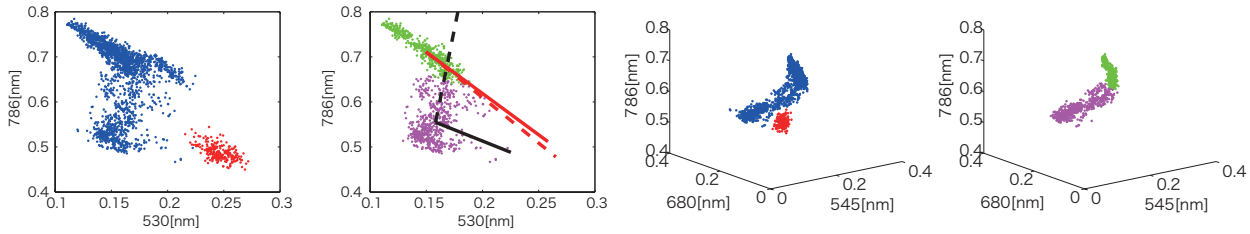


Figure 4. (a) Simulated 2-D normal leaf data and oak wilt data distribution, (b) 2-D data distribution of classified data based on the Gaussian mixture model in manifold space, (c) 3-D normal leaf data and oak wilt data distribution, (b) 3-D data distribution of classified data based on the Gaussian mixture model in manifold space.

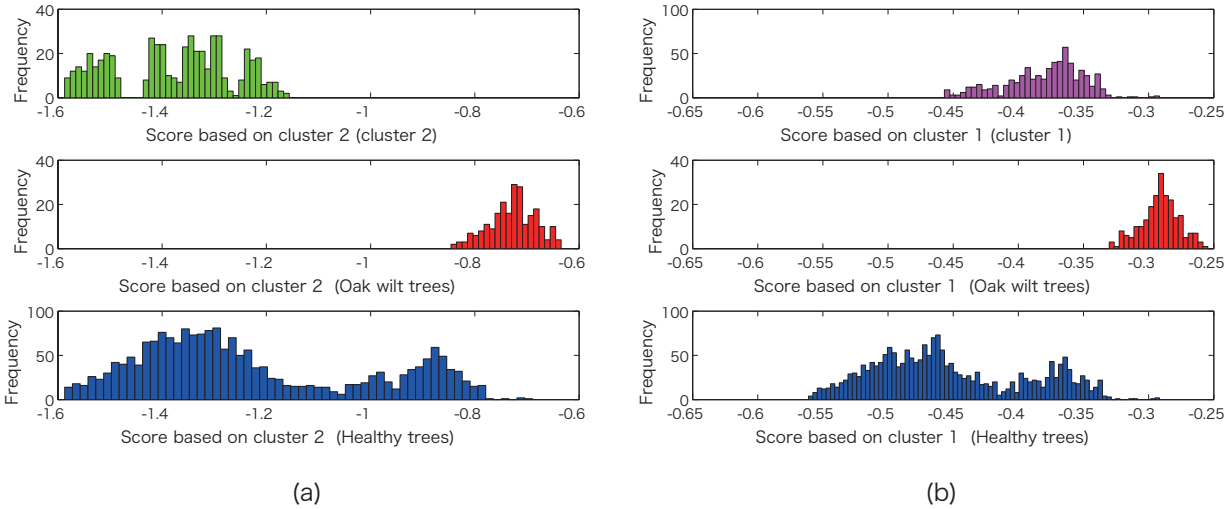


Figure 5. (a) Histograms of scores based on local coordinate vector by Equation (4) with cluster 1, top: Healthy cluster 1 trees, middle: Oak wilt, bottom: Healthy trees, (b) Histograms of scores based on local coordinate vector by Equation (4) with cluster 2, top: Healthy cluster 1 trees, middle: Oak wilt, bottom: Healthy trees.

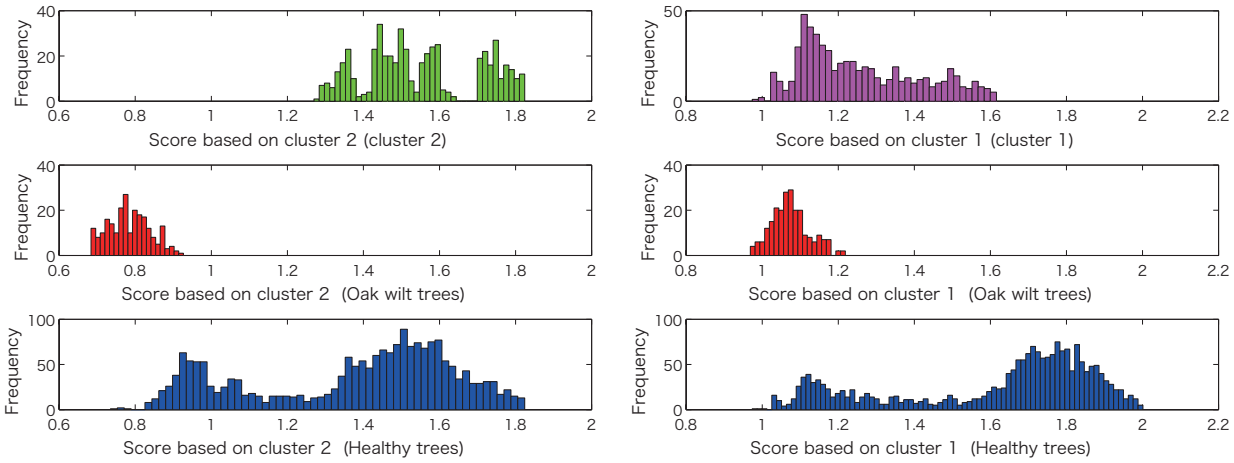


Figure 6. (a) Histograms of scores based on the first principal vector in cluster 1, top: Healthy cluster 1 trees, middle: Oak wilt, bottom: Healthy trees, (b) Histograms of scores based on local coordinate vector based on the first principal vector in cluster 2, top: Healthy cluster 1 trees, middle: Oak wilt, bottom: Healthy trees.

References

Banerjee, A., Burlina, P., Diehl, C., 2006. A support vector method for anomaly detection in hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing, 44 (8), pp.2282-2291.

Gao, B.-c., 1996. NDWI--A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment, 58 (3), pp.257-266.

Heesung, K., Nasrabadi, N.M., 2005. Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing, 43 (2), pp.388-397.

Kamata, N., 2002. Outbreaks of forest defoliating insects in Japan, 1950-2000. Bulletin of Entomological Research, 92 (2), pp.109-117.

McLachlan, G., Peel, D., 2000. Finite Mixture Models. John Wiley & Sons, Inc.

Reed, I.S., Yu, X., 1990. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. IEEE Transactions on Acoustics, Speech and Signal Processing, 38 (10), pp.1760-1770.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., Year. Monitoring vegetation systems in the Great Plains with ERTS. In: Proceedings of the Third ERTS Symposium, Vol.1, pp.317.

Roweis, S.T., Saul, L.K., 2000. Nonlinear Dimensionality Reduction by Locally Linear Embedding. Science, 290 (5500), pp.2323-2326.

Tenenbaum, J.B., Silva, V.d., Langford, J.C., 2000. A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science, 290 (5500), pp.2319-2323.

Uto, K., Takabayashi, Y., Kosugi, Y., Ogata, T., Year. Hyperspectral Analysis of Japanese Oak Wilt to Determine Normalized Wilt Index. In: IEEE International Geoscience and Remote Sensing Symposium, Vol.2, pp.295 -II-298.

Wu, C., 2004. Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery. Remote Sensing of Environment, 93 (4), pp.480-492.

Zhang, Z., Zha, H., 2002. Principal Manifolds and Nonlinear Dimension Reduction via Local Tangent Space Alignment. SIAM Journal of Scientific Computing, 26 (1), pp.313-338.