ISSN: 2321-7782 (Online)
Impact Factor: 6.047
Volume 4, Issue 7, July 2016
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

Nonparametric Weighted Fuzzy Feature Extraction for Hyperspectral Images Classification

R. Priya[1]
Research Sholar, Department of Computer Science
Bharathiar University
Coimbatore, India

Dr. M. Senthil Murugan[2]
Director
A.V.C. College of Engineering
Mannampandal, India

Abstract: Feature extraction plays an essential role in high-dimensional data classification. Many studies show that for extracting hyper spectral image features Nonparametric Weighted Feature Extraction is a powerful tool, but it takes enormous computation time. Later Nonparametric Fuzzy Feature Extraction method is introduced. NWFE uses membership grades estimated by fuzzification procedure of the fuzzy K-nearest neighbor algorithm. NFFE is theoretically suitable for handling linear problems but it is not applicable for the nonlinear problems. A new Nonparametric Weighted Fuzzy Feature Extraction method using NWFE and NFFE is introduced in this paper which produce better classification results than with features extracted from existing methods.

Keywords: Nonparametric Weighted Feature Extraction (NWFE), Nonparametric Fuzzy Feature Extraction (NFFE), Linear Discriminant Analysis (LDA), Nonparametric Discriminant Analysis (NDA), Orthogonal Feature Extraction (ORTH), Removal of Classification Structure (REM), Fuzzy K-Nearest Neighbor (FKNN).

I. Introduction

Feature extraction [1] aims to mitigate the Hughes phenomenon [2,3] or the curse of dimensionality [4] so as to enhance classification performance, which is an important technique for high-dimensional classification problems. To achieve statistical confidence, the Hughes phenomenon describes the ratio of the training samples to the number of features that must be maintained at or above some minimum value [2]. However, it is not necessary to have sufficient training samples to keep the ratio in a high-dimensional classification task. Thus, the ratio can be relatively enlarged without increasing the training samples by using feature extraction to reduce the data dimensionality. Feature extraction uses all the features to construct a transformation matrix that maps the original data to a low-dimensional subspace. As compared with feature selection [5], the main advantages of feature extraction are the information of all the original bands is used and it is easier to use than feature selection for high-dimensional data [6].

In [14] Linear Discriminant Analysis (LDA) is addressed as a useful subspace technique for finding projections to a lower dimensional space with the class separation is maximized as it optimizes the Fisher's score. In addition, LDA has no free parameters to be tuned, and the extracted features are potentially interpretable under linearity assumptions. LDA is mainly focused on image classification and band selection with its good capabilities led to its extensive use and practical exploitation in remote sensing applications.

Nonparametric Discriminant Analysis in [15,16] shows that the between-class scatter matrix is redefined as a new nonparametric between-class scatter matrix and also it addressed two problems of NDA.

Kuo and Landgrebe [7] addressed the nonparametric weighted feature extraction (NWFE) with fully nonparametric within- and between-class scatter matrices. In addition, it solve the singularity problem by including a regularization technique [8].

Some studies [9-11] shows that NWFE is efficient in reducing dimensionality, but it reveals some useful consequences for building a nonparametric feature extraction method.

In [12] a novel nonparametric feature extraction method, called nonparametric fuzzy feature extraction (NFFE), is proposed. NFFE introduces fuzzy membership grades to design its within-class and between-class scatter matrices, and a more general regularization form with the same components in NWFE is adopted to alleviate the singularity problem. Importantly, a theoretical adjustment on features, not applying in LDA, NDA and NWFE, is taken to orthogonalize the features.

NWFFE is introduced in this paper which is best suitable for hyperspectral image classification. It combines the best features of NWFE and NFFE.

The rest of the paper is organized as follows. In Section 2, LDA features are analyzed, NDA and its features are analysed in Section 3, followed by NWFE algorithm in Section 4, NFFE algorithm in Section 5 and NWFFE algorithm in Section 6.

II. Linear Discriminant Analysis

In [14] Linear Discriminant Analysis method is analyzed where it allows us to find a linear transformation matrix G that reduces an original m-dimensional feature vector x to an l-dimensional vector $a = G^T x \in R^l$, where $l<m$. this low dimensional feature space is selected to fulfill a given maximization criterion of separability among class distributions. Particularly, Fisher's criterion is based on maximizing the distance means of the classes and at the same time, minimizing their intra-class variances, $(w) = \arg max_w \left\{\frac{(\mu 2-\mu 1)2}{\sigma 1^2+\sigma 2^2}\right\}$

Since the decision function is $y = w^T$ x, and the means and variances can be trivially defined, once can easily demonstrate that maximizing the Fisher's score is equivalent to maximize the following Rayleigh coefficient with respect to the decision function weight vector w:

$$w^* = \arg\max_w \{J(w)\} = \arg\max_w \left\{\frac{w^T S_b w}{w^T S_w w}\right\} \quad (1)$$

where $S_b = \frac{1}{n}\sum_{i=1}^{K} n_i\,(\mu_i - \mu)(\mu_i - \mu)^{2T}$ is the between-class variance, and $S_w = \frac{1}{n}\sum_{k=1}^{K}\sum_{i\in I_k}(x_i - \mu_k)\,(x_i - \mu_k)^T$ is the within class variance, where μ_k and I_k denote the sample mean and the index set for class k, respectively.

The maximization criterion in (1) can be rewritten as the following maximization problem:

$$G^* = \arg max_G \{trace((G^T S_w G)^{-1} G^T S_b G)\}, \quad (2)$$

which can be demonstrated to be equivalent to :

$$G^* = \arg max_G \{trace((G^T S G)^{-1} G^T S_b G)\}, \quad (3)$$

where $S = S_b + S_w$ is the total scatter matrix, which is the estimate of the common covariance matrix. Note that solving any of the above equivalent problems is only possible if S_w and S are non-singular.

In this context, the scatter matrices can be redefined as :

$$S = H^T H, \quad S = H_w^T H_w, \quad S_b = H_b^T H_b, \quad (4)$$

where

$$H = \frac{1}{\sqrt{n}}(X - \mu_1{}^T) \quad (5)$$

$$H_w = \frac{1}{\sqrt{n}}[X_1 - \mu_1 1_1{}^T, \dots, X_k 1_K{}^T), \quad (6) \text{ and}$$

$$H_b = \frac{1}{\sqrt{n}}[\sqrt{n_1}(\mu_1 - \mu), \dots, \sqrt{n_k}(\mu_k - \mu), \quad (7)$$

where 1 is a column vector of n ones, and 1i is a column vector of ni ones

It is worth nothing that (2) and (3) are generalized eigen-decomposition problems for the scatter matrices $S_w^{-1}S_b$ or $S^{-1}S_b$, respectively. As a consequence, there exist a maximum of K-1 eigenvectors with non-zero eigenvalues, as this is the upper bound on the rank of S_b. Several problems arise from this theoretical property. Firstly, selecting a maximum of K-1 eigenvectors of matrix $S_w^{-1}S_b$ for projecting data and classification may not be always sufficient, because essential information can be lost. Secondly, S or S_w are often singular in small data sets and thus the solution can not be obtained.

III. Non Parametric Discriminant Analysis

NDA redefines the between-class scatter matrix as a new nonparametric between-class scatter matrix according to [15] [16], denoted S_b^{NDA}, as $S_b^{NDA} = P_1 E\left\{\left\{\left(X^{(1)} - M_2\left(X^{(1)}\right)\right)\left(X^{(1)} - M_2\left(X^{(1)}\right)\right)^T\right\}\middle|\omega_1\right\}$

$$+ P_2 E\left\{\left\{\left(X^{(2)} - M_1\left(X^{(2)}\right)\right)\left(X^{(2)} - M_1\left(X^{(2)}\right)\right)^T\right\}\middle|\omega_2\right\} \qquad (8)$$

where $X^{(i)}$ denotes the random variable used to describe the distribution of class i, and $x_l^{(i)}$ denotes the *l*-th outcome of this random variable. $M_j\left(x_l^{(i)}\right) = \frac{i}{k}\sum_{j=1}^{k} x_{jNN}^{(i)}$ is called the local kNN mean, $x_{jNN}^{(i)}$ is the jth nearest neighborhood (NN) from class i (ω_1) to the sample $x_l^{(i)}$. If $K = N_i$, the training sample size of class i, shows that the features extracted by maximizing $tr[(S_w^{NDA})^{-1}\, S_b^{NDA}]$ must be the same as the ones from $tr[(S_w^{DA})^{-1}\, S_b^{DA}]$. Thus, the parametric feature extraction obtained by maximizing $tr[(S_w^{DA})^{-1}\, S_b^{DA}]$ is a special case of feature extraction with the more general nonparametric criterion $tr[(S_w^{NDA})^{-1}\, S_b^{NDA}]$, where the definition of S_w^{NDA} is in (17).

It is appropriate to use method of de-emphasizing samples far from the boundary. To accomplish this, (15) uses a weighting function for each $\left(x_l^{(i)} - M_j\left(x_l^{(i)}\right)\right)$. The value of the weighting function, denoted as w_l, for $x_l^{(i)}$ is defined as

$$w_l^{(i,j)} = \frac{min\left\{d^\alpha\left(x_l^{(i)}, x_{kNN}^{(i)}\right), d^\alpha\left(x_l^{(i)}, x_{kNN}^{(j)}\right)\right\}}{d^\alpha\left(x_l^{(i)}, x_{kNN}^{(i)}\right) + d^\alpha\left(x_l^{(i)}, x_{kNN}^{(j)}\right)} \qquad (9)$$

where α is a control parameter between zero and infinity, and $d\left(x_l^{(i)}, x_{kNN}^{(j)}\right)$ is the Euclidean distance from $x_l^{(i)}$ to its kNN point in class j.

Although the nonparametric version of the within-class matrix was proposed in [15] and [16], the parametric S_w^{DA} was still suggested to optimize $J_{NDA} = tr[(S_w^{DA})^{-1}S_b^{NDA}]$ be used in NDA by the authors.

The difficulties of NDA are (i) Parameters k and α are usually decided by rules of thumb. So the better result usually comes after several trials. (ii) The within-class scatter matrix in NDA is still with a parametric form. When the training set size is small, NDA will have the singularity problem.

IV. Non Parametric Weighted Feature Extraction

The main ideas of NWFE are putting different weights on every sample to compute the "weighted means" and defining new nonparametric between-class and within-class scatter matrices to obtain more than L–1 features [7]. The nonparametric between-class scatter matrix for L classes is defined as

$$S_b^{NW} = \sum_{i=1}^{L} P_i \sum_{\substack{j=1 \\ j\neq i}}^{L} \sum_{l=1}^{N_i} \frac{\lambda_l^{(i,j)}}{n_i}\left(x_l^{(i)} - M_j\left(x_l^{(i)}\right)\right)\left(x_l^{(i)} - M_j\left(x_l^{(i)}\right)\right)^T \qquad (10)$$

where $x_l^{(i)}$ refers to the l-th sample from class i, N_i is training sample size of class i, P_i denotes the prior probability of class i.

The scatter matrix weight $\lambda_l^{(i,j)}$ is a function of $x_l^{(i)}$ and $M_j(x_l^{(i)})$ and defined as:

$$\lambda_l^{(i,j)} = \frac{dist\left(x_l^{(i)}, M_j\left(x_l^{(i)}\right)\right)^{-1}}{\sum_{t=1}^{N_j} dist\left(x_t^{(i)}, M_j\left(x_t^{(i)}\right)\right)^{-1}} \quad (11)$$

where dist**(**a,b) denotes the Euclidean distance from a to b.

If the distance between $x_l^{(i)}$ and $M_j(x_l^{(i)})$ is small then its weight $\lambda_l^{(i,j)}$ will be close to 1; otherwise, $\lambda_l^{(i,j)}$ will be close to 0. The sum of the $\lambda_l^{(i,j)}$ for class i is 1.

$M_j(x_l^{(i)})$ denotes the weighted mean of $x_l^{(i)}$ in class j and defined as

$$M_j(x_l^{(i)}) = \sum_{k=1}^{N_j} w_{lk}^{(i,j)} x_k^{(j)}, \quad (12)$$

where

$$w_{lk}^{(i,j)} \frac{dist\left(x_l^{(i)}, x_k^{(j)}\right)^{-1}}{\sum_{t=1}^{n_j} dist\left(x_l^{(i)}, x_k^{(j)}\right)^{-1}}. \quad (13)$$

The nonparametric within-class scatter matrix is defined as

$$S_w^{NW} = \sum_{i=1}^{L} P_i \sum_{l=1}^{N_i} \frac{\lambda_l^{(i,j)}}{n_i}\left(x_l^{(i)} - M_i\left(x_l^{(i)}\right)\right)\left(x_l^{(i)} - M_i\left(x_l^{(i)}\right)\right)^T \quad (14)$$

The extracted features are the f eigenvectors with largest f eigenvalues of the following matrix:

$S_w^{NW} = 0.5\, S_w^{NW} + 0.5\, diag(S_w^{NW})$, where diag**(**A) means the diagonal parts of matrix A.

The NWFE Algorithm is

Step 1: Compute the distances between each pair of sample points and form the distance matrix.

Step 2: Compute $w_{lk}^{(i,j)}$ using the distance matrix

Step 3: Use $w_{lk}^{(i,j)}$ to compute the weighted means $M_j(x_l^{(i)})$

Step 4: Compute the scatter matrix weight $\left(\lambda_l^{(i,j)}\right)$

Step 5: Compute S_b^{NW} and regularized S_w^{NW}

Step 6: Extract features by using ORTH and REM methods

V. NON PARAMETRIC FUZZY FUTURE EXTRACTION (NFFE)

The ideas behind the nonparametric fuzzy feature extraction (NFFE) are described as follows. In NFFE, we use the membership grades estimated by the fuzzification procedure of the fuzzy K-nearest neighbor (FKNN) algorithm [13] to design its weighting function. The fuzzification procedure of FKNN is defined as

$$\mu_j x_l^{(i)} = \begin{cases} 0.51 + 0.49\, X\, \frac{n_j}{k_1}\; if\ j = i \\ 0.49\, X\, \frac{n_j}{k_1}\; if\ j \neq i \end{cases}$$

where $x_l^{(i)}$ is the lth training sample in the class i, k_1 is a given constant denoting for the k_1-nearest neighbors of $x_l^{(i)}$, n_j is the number of samples that belongs to the class j among the k_1-nearest neighbors of $x_l^{(i)}$ and $\sum_{j=1}^{L} n_j = k_1$.

The membership grades computed by the above equation reflect the information which is employed to design the weighting function of NFFE.

The within-class scatter matrix of NFFE (S_{fw}) is defined as

$$S_{fw} = \sum_{i=1}^{L} P_i \sum_{l=1}^{N_i} a_l^{(i,i)} \left(x_l^{(i)} - M_i(x_l^{(i)})\right)\left(x_l^{(i)} - M_i(x_l^{(i)})\right)^T, \quad (15)$$

where $M_i(x_l^{(i)})$ is the weighted local mean of $x_l^{(i)}$ in class i. $a_l^{(i,i)}$ is the weight with respect to $x_l^{(i)}$.

The between-class scatter matrix of NFFE (S_{fb}) is defined as

$$S_{fb} = \sum_{i=1}^{L} P_i \sum_{\substack{j=1 \\ j\neq i}}^{L} \sum_{l=1}^{N_i} b_l^{(i,j)} \left(x_l^{(i)} - M_j(x_l^{(i)})\right)\left(x_l^{(i)} - M_j(x_l^{(i)})\right)^T, \quad (16)$$

where $M_j(x_l^{(i)}) = \sum_{s=1}^{k_2} \alpha_s^{(j)} x_{sNN}^{(j)}$ and $b_l^{(i,j)}$ are the weighted local mean and the weight of $x_l^{(i)}$ in class j, respectively.

To make NFFE applicable for SSS [Small Sample Size] problems, the regularization of S_{fw} is implemented by including an adaptive parameter for finding a suitable within-class estimator of NFFE. The regularized within-class scatter matrix S_{fw}^R of NFFE is defined as

$$S_{fw}^R = (1-\mu)S_{fw} + \mu diag(S_{fw}), 0 \le \mu \le 1. \quad (17)$$

where μ is the regularization parameter.

To make the eigenvectors v_h's orthonormal with respect to S_{fw}^R to satisfy $A^T S_{fw}^R A = I$, the scale v_h must be adjusted by

$$v_h = \frac{v_h}{\sqrt{v_h^T S_{fw}^R v_h}} \quad (18)$$

NFFE includes three parts: the design of nonparametric scatter matrices, the regularization of the within-class scatter matrix and the adjustment of the extracted features.

The NFFE algorithm is

Step 1: Given a value of k_1 for estimating the membership grades of samples in X.

Step 2: Compute the within-class and between-class weights of each $x_l^{(i)}$ in X, i.e. $a_l^{(i,i)}$ and $b_l^{(i,j)}$.

Step 3: Given a value of k_2 to compute the weighted local means and then compute S_{fw} and S_{fb}.

Step 4: Replace S_{fw} with S_{fw}^R where the optimal value of μ is searched by grid-search and 5-fold CV methods in a grid of 0.05.

Step 5: Select the p eigenvectors of $(S_{fw}^R)^{-1} S_{fb}$, which correspond to the p largest eigenvalues.

Step 6: Adjust each eigenvector v_h and $A = [v_1, \dots, v_p] \in R^{dxp}$

Step 7: Calculate the transformed data $Y = A^T X$.

VI. NON PARAMETRIC WEIGHTED FUZZY FEATURE EXTRACTION METHODS

The idea behind the Nonparametric Weighted Fuzzy Feature Extraction Method is to merge the Nonparametric Weighted Feature Extraction and Nonparametric Fuzzy Feature Extraction which produce better classification results when compared to other approaches.

The NWFFE Algorithm is

Step 1 : Compute the distances between each pair of sample points and form distance matrix.

Step 2 : Compute within-class and between-class weights for each sample using

$$S_w^{NW} = \sum_{i=1}^{L} P_i \sum_{l=1}^{N_i} \frac{\lambda_l^{(i,j)}}{n_i} \left(x_l^{(i)} - M_i\left(x_l^{(i)}\right)\right)\left(x_l^{(i)} - M_i\left(x_l^{(i)}\right)\right)^T \quad (19)$$

and

$$S_b^{NW} = \sum_{i=1}^{L} P_i \sum_{\substack{j=1 \\ j\neq i}}^{L} \sum_{l=1}^{N_i} \frac{\lambda_l^{(i,j)}}{n_i} \left(x_l^{(i)} - M_j\left(x_l^{(i)}\right)\right)\left(x_l^{(i)} - M_j\left(x_l^{(i)}\right)\right)^T \quad (20)$$

Step 3 : Compute weighted means using distance matrix using

$$M_j\left(x_l^{(i)}\right) = \sum_{k=1}^{N_j} w_{lk}^{(i,j)} x_k^{(j)} \quad (21)$$

Step 4: Compute between-class scatter matrix and regularized within-class scatter matrix using $S_{fb} =$

$$\sum_{i=1}^{L} P_i \sum_{\substack{j=1 \\ j\neq i}}^{L} \sum_{l=1}^{N_i} b_l^{(i,j)} \left(x_l^{(i)} - M_j\left(x_l^{(i)}\right)\right)\left(x_l^{(i)} - M_j\left(x_l^{(i)}\right)\right)^T \quad (22)$$

and

$$S_{fw}^{R} = (1-\mu)S_{fw} + \mu diag\left(S_{fw}\right), 0 \leq \mu \leq 1 \quad (23)$$

Step 5 : Extract the features using ORTH and REM.

Experiments were conducted and the results reveal that NWFFE Feature Extraction method produces better results when compared to other methods.

Table 1: Comparison of NWFE, NFFE and Proposed Method

Number of Features (2)		Mean of Accuracies				
		LDA	NDA	NWFE	NFFE	NWFFE
INDIAN PINES DATASET (IPD)	ORTH	0.8739	0.8871	0.9479	0.9784	0.9852
	REM	0.8732	0.8862	0.9481	0.9791	0.9876
WASHINGTON DC MALL (WDC)	ORTH	0.8167	0.8437	0.9218	0.9134	0.9342
	REM	0.8123	0.8324	0.9192	0.9235	0.9428

Table 1 shows the comparison between five feature extraction methods LDA, NDA, NWFE, NFFE and NWFFE. It ensures that when compared to others NWFFE produces better mean of accuracies.

Figure 1 and Figure 2 shows the comparison chart for Indian Pines and Washington DC Mall data sets. It shows that the mean of accuracies of NWFFE is better for using both ORTH and REM when compared to other methods like LDA, NDA, NWFE and NFFE.

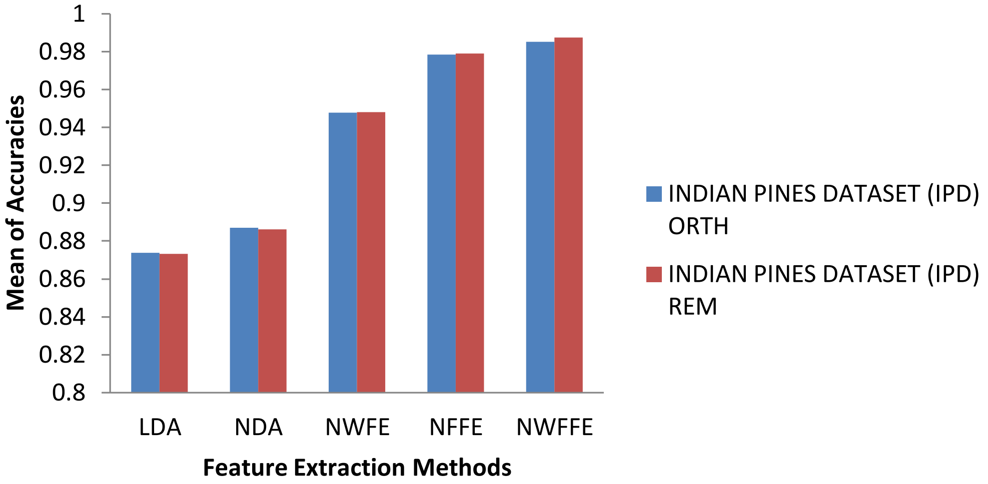

Figure 1: Comparison Chart for Indian Pines Dataset

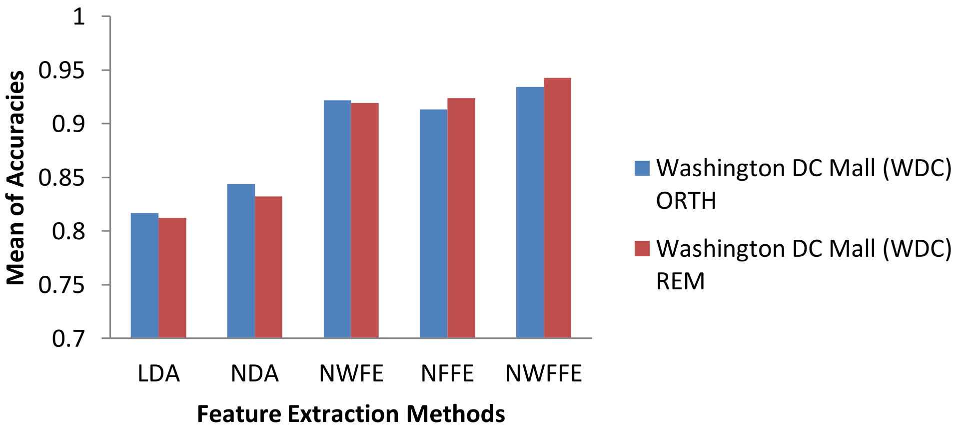

Figure 2: Comparison Chart for Washington DC Mall Dataset

VII. Conclusion

The problems of LDA can be improved by NWFE, but the feature transformation of NWFE is still linear. The most important parts of NWFE are Weighted between- and within-scatter matrices and regularization, by applying one of them cannot get a satisfactory result. NWFE also needs more computation time. NFFE consists of three constituents, that is the construction of sound scatter matrices, the employment of a suitable regularization, and the adjustment of the extracted features. NFFE also recommend adopting the feature adjustment procedure because it is of theoretical and practical importance for the design of such a method. The feature adjustment procedure is of practical importance for reaching better results. NFFE is theoretically suitable for handling linear problems. The NWFFE uses both NWFE and NFFE to extract the features of hyperspectral images and the results shows that it produces better accuracies when compared to LDA, NDA, NWFE and NFFE.

References

1. D. A. Landgrebe, Signal Theory Methods in Multispectral Remote Sensing. Hoboken, NJ: Wiley, 2003.
2. G.. F. Hughes, “On the mean accuracy of statistical pattern recognizers,” IEEE Transactions on Information Theory , vol. 14, no. 1, pp. 55-63, 1968.
3. P. K. Varshney and M. K Arora. Advanced image processing techniques for remotely sensed hyperspectral data , Springer, New York, 2004. International Journal of Fuzzy Systems, Vol. 12, No. 3, September 2010
4. R. O. Duda, P. E. Hart, and D. G. Stock, Pattern Classification , John Wiley & Sons, 2nd, 2001.
5. K. Fukunaga, Introduction to Statistical Pattern Recognition, San Diego, CA:Academic Press, 1990.
6. F. Van der Heiden, R. P. W. Duin, D. de Ridder, and D. M. J. Tax, Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB , Chichester: John Wiley & Sons, 2004.
7. B. C. Kuo and D. A. Landgrebe, “Nonparametric weighted feature extraction for classification,” IEEE Transactions on Geoscience and Remote Sensing , vol. 42, no. 5, pp. 1096-1105, May 2004.

8. H. Friedman, “Regularized discriminant analysis,” Journal of the American Statistical Association, vol. 84, pp.165-175, 1989.
9. P. F. Hsieh, D. S. Wang, and C. W. Hsu, “A linear feature extraction for multiclass classification problems based on class mean and covariance discriminant information, extraction,” IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 28, no. 2, pp. 223-235, Feb. 2006.
10. X. Song, G. Fan, and M. Rao, “Automatic CRP mapping using nonparametric machine learning approaches,” IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 4, pp. 888-897, 2005.
11. J. A. Richards, “Analysis of remotely sensed data: the formative decades and the future,” IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 3, pp. 422-432, 2005.
12. Jinn-Min Yang, Pao-Ta Yu, Bor-Chen Kuo, and Ming-Hsiang Su, “Nonparametric Fuzzy Feature Extraction for Hyperspectral Image Classification”,International Journal of Fuzzy Systems, Vol. 12, No. 3, September 2010.
13. J. M. Keller, M. R. Gray, and J. A. Givens, Jr., “A fuzzy k-nearest neighbor algorithm”, IEEE Transactions on Systems, Man, and Cybernetics , vol. 15, no. 4, pp. 580-58, July/August, 1985.
14. Tatyana V. Bandosa, Lorenzo Bruzzoneb and Gustavo Camps-Vallsa, “Efficient Regularized LDA for Hyperspectral Image Classification”, Proc. of SPIE Vol. 6748.
15. K. Fukunaga and M. Mantock, Nonparametric Discriminant Analysis, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 5, 1983, pp. 671-678.
16. K. Fukunaga, Introduction to Statistical Pattern Recognition, San Diego: Academic Press Inc., 1990.
17. A. Campbell, “Canonical Variate Analysis—A General Model Formulation,” Australian J.Statistics, vol. 26, pp.86-96, 1984.