Comparison of Curve Fitting Method for Hyperspectral Data Classification with Non-linear Based Feature Extraction Methods

M.Sadighi Tehrani
Department of Communication
College of Electrical Engineering
Yadegar-e-Imam Khomeini (RAH)
Shahre- rey branch, Islamic Azad University,
Tehran, Iran
Maryam_st1364@yahoo.com

S.Abolfazl hosseini
Department of Communication
College of Electrical Engineering
Yadegar-e-Imam Khomeini (RAH)
Shahre- rey branch, Islamic Azad University
Tehran, Iran
Abolfazl.hosseini@modares.ac.ir

Batoul zarezardeini
Department of Communication
College of Electrical Engineering
Yadegar-e-Imam Khomeini (RAH)
Shahre- rey branch, Islamic Azad
University, Tehran, Iran
Zare.f1844@yahoo.com

Abstract— **Hyperspectral imagery is one of the most important tools in remote sensing. Increasing the number of bands, lack of training samples, correlation of spectral samples and redundancy of data make conventional image classification methods without reducing the dimensions of the feature vector, not applicable. Dimension reducing as a preprocessing step could be done in two approach: feature selection, and feature extraction. In this paper, recently rational function curve fitting-feature extraction (RFCF-FE) method is analyzed and compared with some nonlinear non-kernel based feature extraction methods such as locally linear embedding (LLE), piecewise constant function approximations (PCFA) and the proposed feature extraction Based on breakpoints (BPB). The maximum likelihood (ML) classification results demonstrate that RFCF-FE provides better classification accuracies compared to competing methods. Also, in this paper, we propose a method for determining the separability of data classes based on specific breakpoints of the spectral response curve (SRC) of pixels.**

Keywords- Hyperspectral images; classification; dimension reduction; feature extraction; curve fitting.

I. INTRODUCTION

Hyperspectral images are represented in the form of 3D data by collecting spatial information in two dimensions x, y and spectral information in the third dimension. These data are generally a combination of images in hundreds of spectral bands with narrow band widths in the electromagnetic spectrum (often in the range of 400 - 2500 nm). Classification is one of the most important issues in hyperspectral imagery. The goal of classification is to associate a proper label to each pixel [1, 2].

Although, it seems that the spectrum of hyperspectral images is complete, but some problems may affect the accuracy of classification. These problems are included: the adverse effects of the earth's atmosphere on the intensity of the pixels in some bands, the great number of bands and the large volume of information, long processing time, low number of training samples relative to the number of image bands and existence of between-bands correlation properties. Therefore, in order to overcome these issues, we should increase the number of training samples or reduce the number of features or use some classification methods which are less sensitive to the number of samples [1, 2].

Feature reduction (either feature selection or extraction approaches) is used to find a transformation that maps data to a lower dimensional space where essential discriminative information is mainly preserved. Feature extraction algorithms could be divided to supervised / non-supervised, linear / non-linear, kernel based / non-kernel based methods [3, 4].

Many approaches such as principal component analysis (PCA) [3], kernel-based principal component analysis (KPCA) [2],linear discriminant analysis (LDA) [1,3],non-parametric weighted feature extraction (NWFE) [5], kernel non-parametric weighted feature extraction (KNWFE) [6], generalized discriminant analysis (GDA) [7], principal curves (PC) [8], locally linear embedding (LLE) [9], piecewise constant function approximations (PCFA) [10] and rational function curve fitting feature extraction (RFCF-FE) [1,11], have been introduced as FE methods. In this paper, we want to compare the RFCF-FE method with non-linear non-kernel based FE methods such as LLE and PCFA.

In Section II, the basic concepts of hyperspectral image classification, feature extraction and general concept of curve fitting are explained. In Section III, we a new proposed method is introduced and compared the RFCF-FE method and other nonlinear feature extraction methods. In section 4, the experiments and results are presented. At the end, the results will be summarized.

II. BASIC CONCEPTS

Image classification is a decision-making process that every image pixels is mapped to a certain class label [12, 13]. Dimension reduction algorithms as a pre-processing step helps to increase the efficiency of the classification [14, 15]. From a perspective, feature extraction methods can be divided into linear approaches such as PCA, LDA, NWFE, discriminant analysis feature extraction(DAFE) and non-linear approaches such as LLE, principal curves and KPCA [8,9,16].

The main problem of the most linear dimension reduction methods is that they cannot keep the local information. However, local information is sufficiently preserved in most of non-linear methods, evaluation of their initial conditions is difficult [17, 18]. These methods are more complicated and difficult to analyze and looking for a locally flat subspace. Non-linear methods either use kernel functions to extend linear models to non-linear models or are inherently non-linear. In kernel-based methods such as KPCA and KNWFE the separation of classes in the kernel space increases with the use of a suitable kernel function for data transformation. In contrast, methods such as RFCF-FE, PCFA, PC and LLE are inherently non-linear. In general, the loss of a part of the spectral information after the dimensions reduction is undeniable.

The RFCF-FE algorithm is a feature extraction method that fit a rational function on each pixel SRC of the hyperspectral data set and considers the coefficients of numerator and denominator polynomials as new feature vectors. This method concentrates on geometrical aspects of the pixel SRC and is supported by this idea that the sequence discipline –ordinance of reflectance coefficients in SRC- includes some information which has not been addressed by many other existing methods that are based on statistical analysis of data. Every SRC is considered as a plot in the form f(λ) [1]. Although, the exact mathematical form of f (λ) is not apparent, the values of N consecutive points are known and can help to fit a curve over these points as an approximation of the original f (λ). Hosseini and Ghassemian represented that it is possible to obtain a mathematical approximation of f (λ) in the form of rational function based on least squares error criterion. Such features also have the ability to compress hyperspectral images [19].

The rational approximation for the function *f*(.) corresponding to the SRC of pixel located at the coordinates (x, y) is:

$$\hat{f}_{(x,y)}\left(\frac{\lambda}{N}\right)=\frac{\sum_{j=0}^{L} c_{j+M+1}\left(\frac{\lambda}{N}\right)^{j}}{1+\sum_{j=1}^{M} c_{j}\left(\frac{\lambda}{N}\right)^{j}} \qquad (1)$$

The overall accuracy (OA) of the ML classifier for the hyperspectral image of the IPS after dimension reduction by RFCF-FE for the 10%, 20%, 30% and 40% volume of training samples is shown in Figure 1. As it has known, this method is not much sensitive to reducing the volume of training samples.

The LLE is a nonlinear feature extraction technique that maps high dimensional data into some low dimensional Euclidean space while preserving local topological structures. LLE includes three steps: "searching for nearest neighbors, calculating reconstructing weights and determining low dimensional embedding" [18]. The PCFA algorithm, which is presented as the fast feature extraction technique, emphasizes at the use of simple averaging of adjacent spectral bands. In fact, the spectral curve is divided into adjacent regions and for each region, the SRC is approximated by a constant value which is calculated through an iteration based procedure. The extracted values are considered as new features [10].

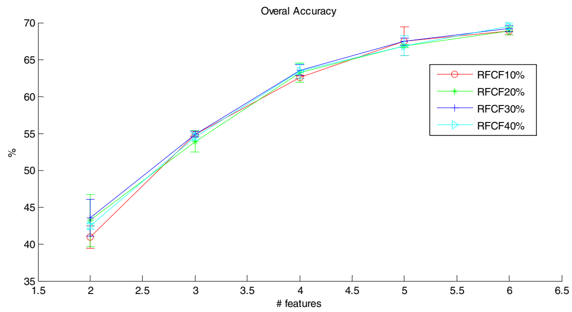

Fig. 1. Overall Accuracy for IPS data set using RFCF-FE with various volumes of training samples

III. The Proposed Method

In this section, a method is proposed to examine the classes separability based on breakpoints of SRCs- the points which the SRC varies suddenly at them. For this purpose, a feature extraction algorithm based on breakpoints (BPB) is proposed.

A. *Feature Extraction Based on Breakpoints*

- Calculate the sum of the absolute values of the difference in consecutive points of SRC (Cumulative storage).
- Determine threshold level by division of the sum to the number of extractive features (*nf*).
- Compare the sum of the absolute differences between the current point and the previous points until it is equal or greater than the threshold.
- Determine the breakpoint and repeat the previous step for the remaining points.
- Averaging the intensities of points between the breakpoints and determining the new features.

B. *Determining the Class Separability Based on Breakpoints of SRCs* **:**

After determining the breakpoints of SRCs for each pixel, we will perform their statistical analysis for a number of different features. By plotting the histogram, the frequency distribution of breakpoints for the training data of each class can be verified. Histogram of the frequency distribution of each breakpoint for all classes presents that these features can be used to determine the separability of the pairwise classes. Therefore, for a special number of features, and for some combination of two classes, there may be some breakpoints (for example, the third breakpoint of five breakpoints) which its range for each of that two classes is completely separate, and they have no overlap.

IV. Experimental Results

A. *Hyperspectral Data Sets*

The Indian Pine (IPS) data set is a mixed forest/agricultural image captured by the Air-borne Visible /Infrared Imaging Spectrometer (AVIRIS). This 145×145 pixels size image includes 220 spectral bands in the wavelength range of 400 to 2500 nm with the spectral resolution of 10 nm. The spatial resolution of this data set is 20m. Twenty water absorption bands were removed and N=200 bands were left. The class map of data includes 16 different classes (see Fig 2) [1].

The other data set, Washington DC Mall belongs to an urban area, is a Hyperspectral Digital Imagery Collection Experiment (HYDICE) airborne hyperspectral data flight line over the Washington, DC-Mall. 210 bands are collected in the range of 400-2400 nm. 191 channels are preserved after discarding some water absorption channels. The data set contains six land cover classes: roof, road, grass, tree, water and shadow (see Fig. 3) [20].

B. *Experiments and Results*

According to the definition of separable classes, it was observed that some image classes, for a number of specific features, have some completely separating breakpoints. This separability is usually greater for classes with fewer members. In the DC-MALL image, classes 3 (grass) and 6 (shadow), for the different feature number are often separated. Fig. 4 shows that when each SRC in the DCMALL image is divided into five parts based on breakpoints, in the third breakpoint, the classes 3 and 6 (grass and shadow) are completely separable. Also, Fig. 5 shows that when each SRC is divided into 7 parts based on breakpoints, the same classes are also completely separable for the fifth breakpoint. The superiority of the RFCF-FE method versus some well-known linear feature reduction techniques such as PCA, LDA, and NWFE and some nonlinear kernel based methods like KPCA, GDA, and MMP has been investigated by [1, 19]. In this article some Non-linear non-kernel based methods have been compared to RFCF-FE.

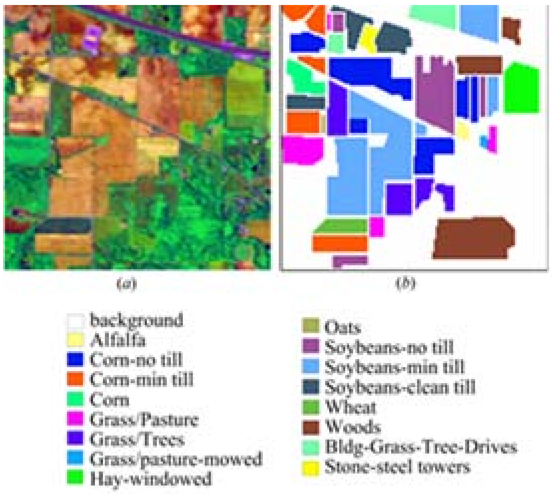

Fig. 2. a) False color image of IPS b) its ground truth map [1,22]

Fig. 3. False color image of DCMALL [20,21]

The RFCF, LLE, PCFA, and BPB algorithms have been implemented to reduce the dimension on hyperspectral data. Fig. 6 shows the IPS accuracy diagram for 10% volume of training samples. The classifier is ML and experiments are performed for 4 to 10 features. Fig. 7 shows the results of this experiment for RFCF, PCFA and BPB methods on the DC-MALL image. According to the results, by increasing the number of features, RFCF-FE method has the best accuracy compared to the competing methods. All measures (average accuracy (AA), average validity (AV), overall accuracy (OA), and kappa statistic) have been improved by RFCF-FE algorithm in comparison to other FE techniques. This improvement for IPS data set is more than that of DCMALL.

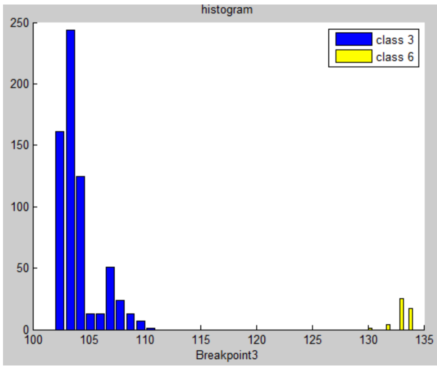

Fig. 4. Classes 3 and 6 (Grass and Shadow) of DCMALL at the third breakpoint for $nf = 5$

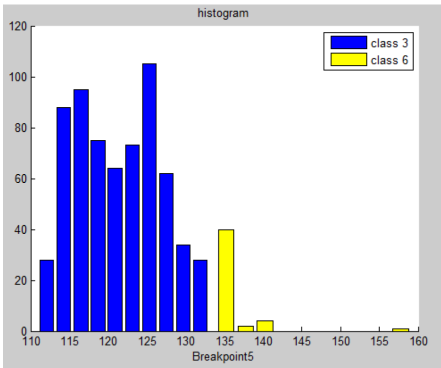

Fig. 5. Classes 3 and 6 (Grass and Shadow) of DCMALL at the fifth breakpoint for $nf = 7$

V. CONCLUSION

In this paper, the RFCF-FE method was investigated and compared with nonlinear non-kernel based feature extraction methods. RFCF-FE is a well performance technique to reduce features using a least-squares error criterion. It obtains a mathematical approximation of the SRCs of pixels in the form of a rational function with polynomials in the numerator and denominator. The coefficients of these polynomials are used as new features. Results after using the ML classifier showed that the RFCF-FE method is superior to competing methods and increases the classification accuracy. Also, the PCFA method is in the second level. The PCFA method presents an appropriate approximation of SRC of each pixel and this is capable to increase the accuracy of the classification to the desired level. As the number of extracted features increases, the accuracy is improved.

REFERENCES

[1] S.A. Hosseini and H. Ghassemian, "Hyperspectral Data Feature Extraction Using Rational Function Curve Fitting", *International Journal of Pattern Recognition and Artificial Intelligence,* vol. 30, p. 1650001, 2016.

[2] M.Imani and H.Ghassemian "Feature extraction of hyperspectral images using boundary semi-labled samples and hybrid criterion", journal of AI and Data Mining ,Vol.5,No1,2017.

[3] W.Liao, "Feature Extraction and Classification for Hyperspectral Remote Sensing Images", dissertation of technology and engineering , Developed by Ghent University Library ,2012

[4] S.P.Lodha and S.M.Kamlapur, "Dimensionality Reduction Techniques for Hyperspectral Images",International Journal of Application or Innovation in Engineering &Management (IJAIEM),Volume 3,Issue 10,2014.

[5] Kuo,B.-C.and Landgrebe, D.A,"Nonparametric Weighted Feature Extraction for Classification Geoscience and Remote Sensing",IEEE Transactions on, 2004.

[6] B.C.Kuo,C.H.Li and J.M.Yang,"Kernel Nonparametric Weighted Feature Extraction for Hyperspectral Image Classification ",IEEE Transactions geoscience and remote sensing ,Vol.47,NO.4,2009.

[7] Baudar,G and Anouar.F,"Generalized Discriminant Analysis Using a Kernel Approach",Neural computation,2000.

[8] K.chang and J.ghosh,"Principal Curves for Nonlinear Feature Extraction and Classification",department of Electrical and Computer Engineering ,1998.

[9] L.k.Saul and S.T.Roweis,"Nonlinear Dimensionality Reduction by Locally Linear Embedding ",Science,Vol.290,Issue 5500,pp,2323,2000.

[10] A.C.Jensen and A.S.Solberg,"Fast Hyperspectral Feature Reduction Using Piecewise Constant Function Sapproximation ",IEEE geoscience and remote sensing letters,Vol.4,2007.

[11] S.A.Hosseini and M.Beitollahi,"Using curve fitting for spectral reflectance curve intervals in order to hyperspectral data compression", in communication systems,Networks and Digital Signal Processing (CSNDSP), 10th international Symposium on 2016.IEEE.2016.

[12] R.Ablin and C.H.Sulochana,"A Survey of Hyperspectral Images Classification in Remote Sensing",International Journal of advanced Research in Computer and Communication Engineering,Vol.2,Issue 8, 2013.

[13] A.b.Santos,C.S.Celes,A.Araujo and D.Menotti,"Feature Selection for Classification of Remote Sensing Hyperspectral Images:A Filter approach using Genetic Algorithm and Cluster Validity ",International Conference on Image Processing,Computer Vision and Pattern Recognition (IPCV'12),Vol.2,2016.

[14] R.Sinduja,S.Chidambaram and A.Sumathi,"Analysis of Dimensionality reduction Techniques for Hyperspectral Image Classification", International Journal of Engineering Trends and Technology (IJETT),Vol.21,NO.2,2015.

[15] X.Jia,B.C.Kuo and M.M.Crawford,"Feature Mining for Hyperspectral ImageClassification",Proceedings of the IEEE,Vol.101,2012.

[16] K.Arunasakthi,L.KamatchiPriya,"A Review on Linear and Non-linear Dimensionality Reduction techniques", Machine Learning and Applications : An International Journal (MLAIJ) Vol.1,No.1,2014.

[17] Sh.Wang and Ch.Wang,"Research on dimension reduction method for hyperspectral remote sensing image based on global mixture coordination factor analysis",International Archives of the Photogrammetry,Remote Sensing and Spatial Information Sciences,Vol.XL-7/W4,2015.

[18] T.Han andD.G.Goodenough, "Nonlinear Feature extraction of hyperspectral Data Based on Locally Linear Embedding (LLE)",IEEE International ,Geoscience and Remote Sensing Symposium (IGARSS), Proceedings,2005.

[19] S.A.Hosseini and h.ghassemian, "Rational function approximation for feature reduction in hyperspectral data",Remote Sensing Lettres Journal,Vol.7,Issue 2, 2015.

[20] G.C-Valls, L.Bruzzane,"Kernel Methods for Remote Sensing Data Analysis",Technology &Engineering Book,2009.

[21] https://engineering.purdue.edu/~biehl/multispec/hyperspectral.html

[22] http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes

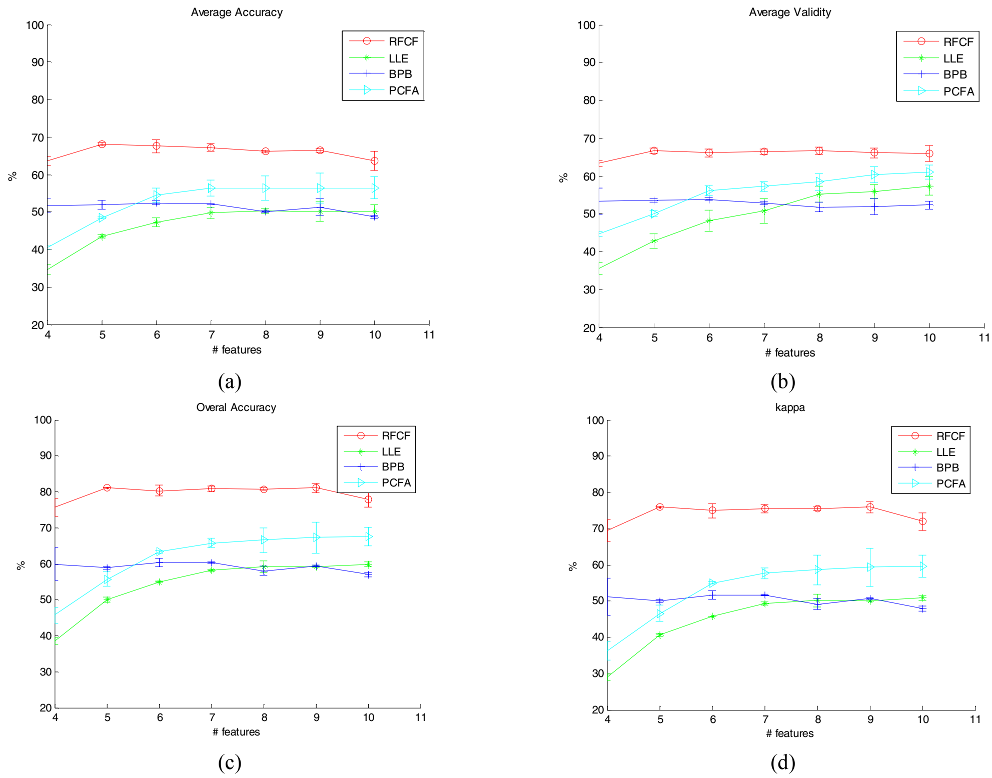

Fig. 6. Classification accuracy measures for IPS data set using different FE methods(RFCF, LLE, PCFA, BPB) and various number of features,(a) average accuracy (b) average validity (c) overal accuracy (d) kappa statistic

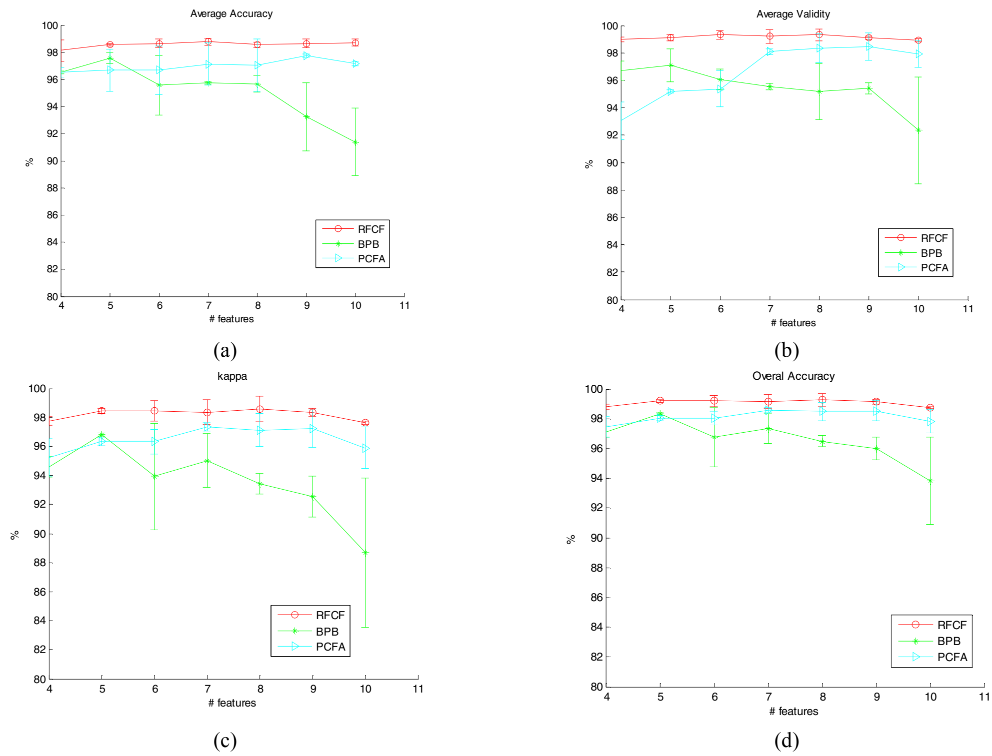

Fig. 7. Classification accuracy measures for DCMALL data set using different FE methods(RFCF, PCFA, BPB) and various number of features,(a) average accuracy (b) average validity (c) overal accuracy (d) kappa statistic