IJCSN International Journal of Computer Science and Network, Volume 4, Issue 4, August 2015
ISSN (Online) : 2277-5420 www.IJCSN.org
Impact Factor: 0.417

Hyperspectral Image Classification Using Harmonic Analysis Integrated with BFO Optimized SVM

[1] Bhanupriya Gaikwad, [2] Vijaya Musande

[1] Department of CSE, MGM's JNEC (Dr. BAMU University), Aurangabad, Maharashtra, India

[2] Department of CSE, MGM's JNEC (Dr. BAMU University), Aurangabad, Maharashtra, India

Abstract - The classification of hyperspectral images is more challenging because of the very high dimensionality of the pixels and the small number of labelled examples typically available for learning. These peculiarities lead to particular signal processing problems, mainly characterized by indetermination and complex manifolds. In our project a new novel method has been introduced that is Harmonic Analysis based classification such as HA-BFO-SVM approach. This new approach accurately classifies the cluster band with respect to their amplitude and phase. Harmonic Analysis (HA) is introduced to extract the feature from hyperspectral image. Amplitude and phase a feature has been obtained by derived HA. Then select best feature among extracted feature by Bacterial Foraging Optimization (BFO). Finally, classify the respective band with related cluster which is performed with the help of Support Vector Machine (SVM). This classifier accurately classifies the band to respective cluster form. In prior work, instead of HA, used MNF, PCA, and ICA could extract features and also classification has been performed by BFO-SVM instead of using PSO-SVM, CV-SVM and GA-SVM.

Keywords - *Harmonic analysis (HA); hyper spectral image classification (HSI); Bacterial Foraging optimization (BFO); Support vector machine (SVM).*

1. Introduction

Hyperspectral imaging, like other spectral imaging, collects and processes information from across the electromagnetic spectrum [1]. The goal of hyperspectral imaging is to obtain the spectrum for each pixel in the image of a scene, with the purpose of finding objects, identifying materials, or detecting processes [2][[3][4]. Much as the human eye sees Visible light in three bands (red, green, and blue), spectral imaging divides the spectrum into many more bands [5][6][7]. This technique of dividing images into bands can be extended beyond the visible. In hyperspectral imaging, the recorded spectra have fine wavelength resolution and cover a wide range of wavelengths [8][9]. Hyper spectral RS information processing are more complex due to curse of dimensionality, limited availability of training samples, the overlapped spectra, the redundancy of spectral information and huge processing data pose Critical challenges to supervised classification of hyper spectral images since theoretical and practical problems arise by the increase of dimensionality [10][11]. Hence it is difficult to apply traditional classifiers to hyper spectral data classification.

To address the ill-posed classification problems, many advanced methods in statistic learning, pattern recognition, computer vision, machine learning, and signal processing fields have been proposed in recent years, among which support vector machine (SVM) is one of the most popular methods [12][13][14]. Usually, SVM outperforms the traditional classifiers, such as maximum likelihood classifier (MLC), artificial neural network (ANN), and spectral angle mapper (SAM), which tends to be more robust in high-dimensional signal processing [15][16]. SVM has such advantages as good generalization capacity , less limit in training set size and more robust to noise, which greatly flourish the applications of SVM to hyperspectral image classification [17][18][19]. However, SVM still has some disadvantages due to the dubious determination of the parameters, such as kernel parameter and penalty parameter [20][21]. Sometimes, SVM often leads to over-fitting problem from optimizing the parameters to model selection [22][23]. Therefore, when applying SVM to solve classification problem, the

parameters including the penalty parameter and the kernel parameter directly affect the final classification accuracy, making the parameters optimization an important issue.

Bacterial foraging optimization algorithms (BFOA) are inspired through the social foraging behaviour of Escherichia coli. It is an algorithmic approximation technique mimicking bacterial colony growth. The four principal mechanisms observed in bacterial are chemotaxis, swarming, reproduction, and elimination-dispersal [26]. Bacterial Foraging Optimization (BFO) algorithms are a new class of biologically encouraged stochastic global search technique based on mimicking the foraging behaviour of E. coli bacterial. This method is used in locating, handling, and ingesting the food. During foraging, a bacterium can exhibit two different actions: tumbling or swimming. The tumbling action modifies the orientation of the bacterium. During swimming means the chemotaxis step, the bacterium will move in its current direction. Chemotaxis movement is continued until a bacterium goes in the direction of positive-nutrient gradient. After a certain number of complete swims, the best half of the population undergoes the reproduction and eliminating the rest of the population. In order to escape local optima, an elimination dispersion event is carried out where some bacterial are liquidated at random by a very small probability and the new replacements are initialized at random locations of the search space. Depends on the fitness value we can choose the best particle [27].

Harmonic analysis (HA) technique has been widely used for power pollution elimination. In remote sensing field, HA is mainly used for analysing the time series data set. Designed for feature extraction, Harmonic analysis on a single pixel, a smooth curve in frequency domain represented by amplitude, phase, and the remainder can be obtained, which is considered to be more functional and discriminative in classification. Different of the traditional feature extraction techniques, a can, the best capture the functional relationships between spectral bands taking into account the adjacent bands to extract the frequency features. HA can also be used as a curve fitting technique that could extract phonological rule within a year from noisy NDVI time series data set. These applications hint at the potential use of HA for hyperspectral image classification. Nevertheless, there are some other optimizing strategies, such as cross validation (CV) [28] and genetic algorithm (GA) [29], which are all used for comparison in our work.

In Existing System feature extraction have been performed via the MNF [23], PCA [24], and ICA method [25]. In feature selection part carried out by PSO, GA and CV algorithms[26]. The minimum noise fraction (MNF) transform is used to determine the inherent dimensionality of image data to segregate noise in the data and to reduce the computer requirements in subsequent processing [24][25]. The first transformation applied to the data decorrelates and recalls the noise in the data based on an estimated noise covariance matrix calculated in the scene prior to the MNF transformation [26]. This results in transformed data in which the noises have unit variance and no band-to-band correlations [27]. Then standard PC transformations are applied to the noise-whitened data.

Principal component analysis (PCA) are a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linear uncorrelated variables called principal components [28]. The number of principal components are less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance. In signal processing, independent component analysis (ICA) are computer methods in separating a multivariate signal into additive subcomponents [29]. This is done by assuming that the subcomponents are non-Gaussian signals and that they are statistically independent of each other. ICA is special case of blind source separation. Common example applications are the "cocktail party problem" of listening in on one person's speech in a noisy room. A genetic algorithm (GA) is a search heuristic that mimics the process of natural selection. This heuristic are routinely used to generate useful solutions to optimization and search problems [30].

2. Proposed Methodology

HA is first proposed to transform the pixels from spectral domain into frequency domain expressed by amplitude, phase and residual, yielding more functional and discriminative features for classification purpose. The Reconstructed Image is adopted with BFO to optimize the penalty parameter C and the kernel parameter for SVM. Finally, the extracted features are classified with the optimized model.
The main novelty and contributions of our work consist of twofold.

I. By acting HA on a single pixel, a smooth curve in frequency domain represented by amplitude, phase, and remainder can be obtained, which is considered to be more functional and discriminative for classification.

II. Different from the traditional feature extraction techniques, HA can better capture the functional relationships between spectral bands taking into account the adjacent bands to extract the frequency features.

The non-zero weights in the SVM are not associated with examples close to the decision boundary, but rather appear to represent 'prototypical' examples of classes. We term these examples 'relevance' vectors, in difference to the principle of automatic relevance determination (ARD) which motivates the presented approach.

2.1 Advantages

a. The proposed feature extraction method shows robust performance when dealing with noisy data set.

b. The proposed feature extraction method (HA) can capture more functional information that are useful for classification purpose

c. The BFO optimization method has advantages such as local minima, randomness, direction of movement, attraction/repelling, swarming and so on.

d. Standard formulation as a probabilistic generalized linear model. This implies that it can be extended to the multiple-class case in a straightforward and principled manner, without the need to train and heuristically combine multiple dichotomous classifiers as is standard practice for the SVM.

3. Module Descriptions

A. Preprocessing the hyperspectral Data

B. HA-based Feature Extraction

C. BFO Optimization

D. SVM Classification

A. Preprocessing

- Input image need to read from database.
- The input image and filtering part to process the MATLAB domain, the process is mainly focused by image processing technique.
- We use the MATLAB in preprocessing steps and to modify the reverse process.
- The middle filters architecture process to perform the Xilinx software using VHDL language. Because the filter architecture process and the internal architecture mainly based upon the bit-wise process.

B. HA-based Feature Extraction

HA is designed to express the time series using overlapped sine and cosine waves (harmonics), in which a signal in time domain can be transformed into the frequency domain represented by multiple overlapped sine and cosine waves with different frequencies. Generally, any periodic waves can be expanded into Fourier series forms. HA technique to hyperspectral image features extraction. Although the traditional Fourier analysis (FA), which is a two-dimensional transformation, can be used to transform hyperspectral image into a range of frequency images associated to each band in frequent domain, the reflected frequency changes of every pixel are based on specific band.

C. BFO Optimization

Bacterial foraging optimization algorithm (BFOA) is inspired by the social foraging behaviour of Escherichia coli. It is an algorithmic approximation technique mimicking bacteria colony growth. The four principal mechanisms observed in bacteria are chemotaxis, swarming, reproduction, and elimination-dispersal.

Bacteria Foraging Optimization (BFO) algorithm is a new class of biologically encouraged stochastic global search technique based on mimicking the foraging behaviour of E. coli bacteria. This method is used for locating, handling, and ingesting the food. During foraging, a bacterium can exhibit two different actions: tumbling or swimming. The tumble action modifies the orientation of the bacterium. During swimming means the chemotaxis step, the bacterium will move in its current direction. Chemotaxis movement is continued until a bacterium goes in the direction of positive-nutrient gradient. After a certain number of complete swims, the best half of the population undergoes the reproduction and eliminating the rest of the population. In order to escape local optima, an elimination dispersion event is carried out where some bacteria are liquidated at random with a very small probability and the new replacements are initialized at random locations of the search space.

The Chemotaxis behaviour of E. coli bacteria considering the foraging behaviour of E. coli, it has a common type of bacteria with a diameter of 1 and a length of about 2 and which under appropriate circumstances can reproduce in 20 min. It is the ability to move comes from a set of up to six rigid 100–200 rps spinning flagella, each driven by a biological motor. The E. coli bacterium alternates between running (at 10–20 sec, but they cannot swim straight) and tumbling (changing the direction). When the flagellas are rotate in clockwise, they operate as propellers and hence an E.Coli may be run or tumble.

The Chemotaxis Actions are below:

- If in neutral medium, the alternate tumbles and the runs ⇒ search.
- If swimming up in a nutrient of the gradient or out of the noxious substances, swim longer (climb up nutrient gradient or down noxious gradient) ⇒ seek increasingly favourable environments.
- If swimming down the nutrient of the gradient (or up noxious substance gradient), then search ⇒ to avoid unfavorable environments.

Bacterial Forging Optimization Algorithm which is simple and easy to implement in order to optimize the cost for processing Client Job. BFO has been successfully applied in many areas of power system. This algorithm is very effective in giving quality solutions. It is motivated by the natural selection which contains tends to eliminate the animals with poor foraging strategies and the favor those having successful foraging strategies.

The BFO algorithm can be explained by four processes namely, chemotaxis, swarming, reproduction, and elimination-dispersal. Below we briefly describe each of these processes.

Steps in BFO Algorithms:

I. Chemotaxis:

This process simulates the swimming and tumbling movements of an E. coli cell by a set of rigid flagella. An E. coli bacterium can move in two different ways. It can swim for a period of time in the same direction or it may tumble, and alternate between these two modes of operation for the entire lifetime. This alternation between the two modes enables the bacterium to move in random directions and search for nutrients.

II. Swarming:

It is always desired that when any one of the bacteria reaches the better location, try to attract other bacteria so that they reach the desired place more rapidly. The effect of swarming is to make the bacteria congregate into groups and move as concentric patterns with high bacterial density.

III. Reproduction:

The least healthy bacteria eventually die while each of the healthier bacteria (each with the lower cost function) asexually split into two bacteria, which are then placed in the same location. Thus, the population size after reproduction is maintained constant.

IV. Elimination and Dispersal:

Gradual or sudden changes in the location where a bacterium population lives may occur due to noxious substance, the temperature rises abruptly in the area or some other influence. Events can kill or disperse all the bacteria in a region. This reduces the chances of convergence at local optima location.

3.1 Classification Techniques

- Based on the training samples used in HSI
- Based on the pixel used in HSI

Based on the training samples used in HSI it will be classified as;

- Supervised classification
- Unsupervised classification

In *Supervised classification*, the identity and location of some of the land cover types such as urban, wetland, forest etc., are known as priori through a combination of field works and top sheets. The analyst attempts to locate specific sites in the remotely sensed data that represents homogeneous examples of these land cover types. These areas are commonly referred as training sites because the spectral characteristics of these known areas are used to _train‘ the classification algorithms in eventual land cover mapping of reminder of the image. Multivariate statistical parameters are calculated in each training site. Every pixel both within and outside these training sites are then evaluated and assigned to a class of which it have the highest likelihood of being a member.

In an *Unsupervised classification*, the identities of land cover types have to be specified as classes within a scene are not generally known as priori because ground truth are lacking or surface features within the scene are not well defined. The computers are required to group pixel data into different spectral classes according to some statistically determined criteria. The comparisons in the medical area are the labelling of cells based on their shape, size, colour and texture, which act as features. These methods are also useful in The MRI images.

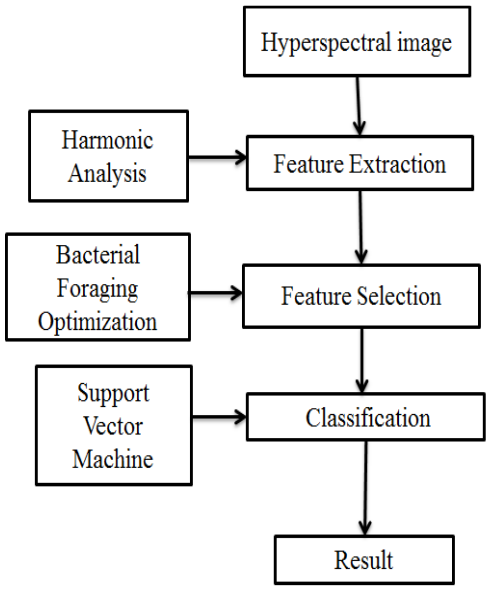

Fig.1.Flow Chart of Module Descriptions

The subsequent Fig.1 shows the flowchart for module descriptions. First the hyper spectral image as given as an input to the harmonic analysis function. Through HA frequency features are extracted. Furthermore the output of HA are given to the Bacterial foraging optimization (BFO) in this optimization technique are used to choose best fitness value among the circumstances. To end with the selected features are given to the support vector machine and it classifies the hyper spectral image. Fig.1 illustrates the proposed HA technique for hyperspectral image representing and feature extraction scheme. By extracting and analysing the features from different spectral domains, the hyperspectral image is transformed into a group of Features of HA, which can maintain the spatial characteristics of hyperspectral data. Precisely, for single pixel, the pixel vector can be expressed by a smooth and complicated curve constructed by a series of sine and cosine waves, during which the Spatial information is reserved. The physical significance of this transformation lies in the ability of reflecting the average energy of single pixel for each band, the energy changes between different bands and the band indices where the amplitude Occurs.

D. SVM Classification

Support vector machines are supervised learning models with associated learning algorithms that analyse data and recognize patterns, used for classification and regression analysis. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

I. Data setup: our dataset contains three classes, each N samples.
II. SVM with linear kernel (-t 0). We want to find the best parameter value C using 2-fold cross validation (meaning use 1/2 data to train, the other 1/2 to test).
III. After finding the best parameter value for C, we train the entire data again using this parameter value.
IV. Plot support vectors
V. Plot decision area

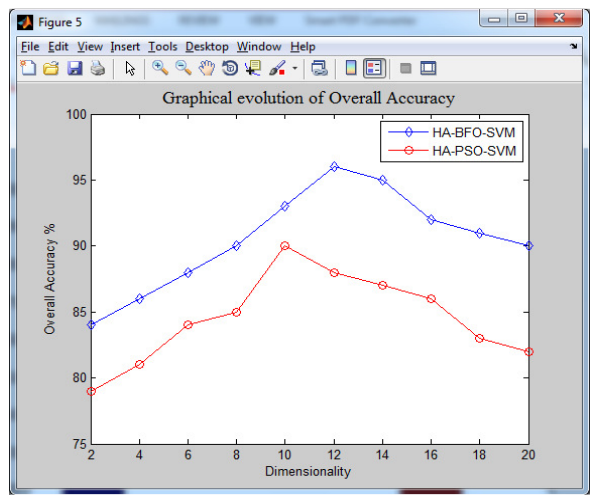

Fig.2.Comparision graph Between PSO-SVM and BFO-SVM

SVM maps input vectors to a higher dimensional vector space where an optimal hyper plane is constructed. Among the many hyper planes available, there is only one hyper plane that maximizes the distance between itself and the nearest data vectors of each category. This hyper plane which maximizes the margin is called the optimal separating hyper plane and the margin is defined as the sum of distances of the hyper plane to the closest training vectors of each category.

Expression for hyper plane:

$$\mathbf{w.x+b = 0}$$

x – Set of training vectors
w – Vectors perpendicular to the separating hyper plane

b – Offset parameter which allows the increase of the margin

Classification by n-class SVM:

This defines a grouping of all the classes in two disjoint groups of classes. This grouping is then used to train a SVM classifier in the root node of the decision tree, using the samples of the first group as positive examples and the samples of the second group as negative examples. The classes from the first clustering group are being assigned to the first (left) subtree, while the classes of the second clustering group are being assigned to the (right) second subtree. The process continues recursively until there is only one class per group which defines a leaf in the decision tree.

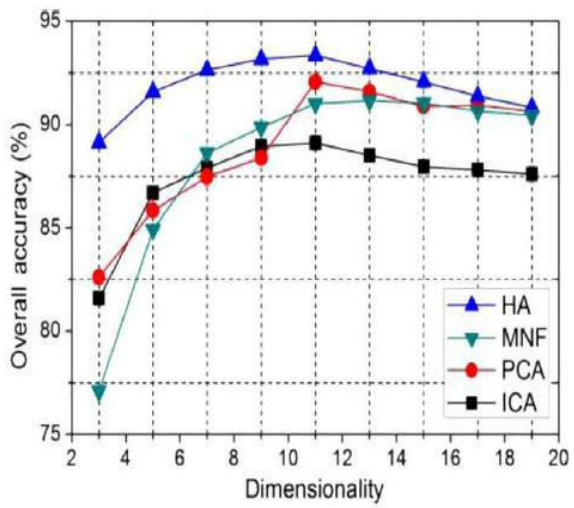

Fig.3. PSO-SVM

Fig.2, Fig.3 and Fig.4 shows the comparison graph for PSO-BFOa, PSO-SVM and GA-SVM. The proposed HA technique based classifiers obtain the best classification accuracies under different dimensionalities. Another observation is that, when the dimensionality is larger than 11, all the classification accuracies decrease fast, especially for GA-SVM. Compared with GA-SVM high accuracy is achieved in PSO-SVM.

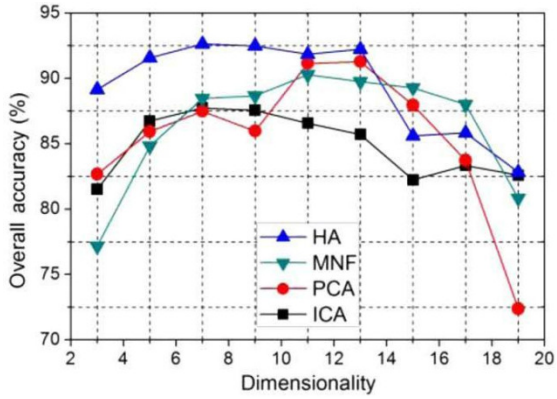

Fig.4. GA-SVM

4. Conclusion

Developed a novel HA-based feature extraction method, exploited a BFO optimized SVM classification scheme, and evaluated the performance of the proposed HA-BFO-SVM classification scheme with respect to different combinations of feature extraction and parameter optimization methods. The proposed method leads to improved performance, and the BFO optimized SVM presents a good trade-off between accuracy and computational time. Specifically, we adapt the HA technique for hyperspectral image analysis and transform the spectral domain into frequency domain represented by amplitude and phase features that are experimentally proofed sensitive and discriminative for classification purpose. That is due to the fact that HA takes adjacent spectral band into account to generate these features, capturing more functional information between bands. Further experiments with additional scenes and comparison methods should be conducted in future. Furthermore, we also envisage the following future perspectives for the development of the presented work.

References

[1] J. A. Richards and X. Jia, Remote Sensing Digital Image Analysis: An Introduction, 4th ed. New York, NY, USA: Springer-Verlag, 2006.

[2] J. Pontius, M. Martin, L. Plourde, and R. Hallett, “Ash decline assessment in emerald ash borer-infested regions: A test of tree-level, hyperspectral technologies,” Remote Sens. Environ., vol. 112, no. 5, pp. 2665–2676, 2008.

[3] U. Amato, A. Antoniadis, M. F. Carfora, P. Colandrea, V. Cuomo, M. Franzese et al., “Statistical classification for assessing PRISMA hyperspectral potential for agricultural land use,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 6, no. 2, pp. 615–625, Apr. 2013.

[4] G. Camps-Valls, L. Gómez-Choval, J. Calpe-Maravilla, J. D. Martin- Guerrero, E. Soria-Olivas, L.

Alonso-Chorda et al., "Robust support vector method for hyperspectral data classification and knowledge discovery," IEEE Trans. Geosci. Remote Sens., vol. 42, no. 7, pp. 1530–1542, Jul. 2004.

[5] Y. Bazi and F. Melgani, "Toward an optimal SVM classification system for hyperspectral remote sensing images," IEEE Trans. Geosci. Remote Sens., vol. 44, no. 11, pp. 3374–3385, Nov. 2006.

[6] G. Camps-Valls, D. Tuia, L. Bruzzone et al., "Advances in hyperspectral image classification," IEEE Signal Process. Mag., vol. 31, no. 1, pp. 45–54, Jan. 2014.

[7] P. Du, K. Tan, and J. Xia, "A novel binary tree support vector machine for hyperspectral remote sensing image classification," Optics Commun., vol. 285, pp. 3054–3060, Jun. 2012.

[8] G. C. Cawley and N. L. C. Talbot, "Over-fitting in model selection and subsequent selection bias in performance evaluation," J. Mach. Learn. Res., vol. 11, pp. 2079–2107, Jul. 2010.

[9] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," Remote Sens. Environ., vol. 113, pp. S110–S122, 2009.

[10] F.Melgani and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," IEEE Trans. Inf. Technol. Biomed., vol. 12, no. 5, pp. 667–677, Sep. 2008.

[11] M. Chi, R. Feng, and L. Bruzzone, "Classification of hyperspectral remotesensing data with primal SVM for small-sized training dataset problem," Adv. Space Res., vol. 41, no. 11, pp. 1793–1799, Mar. 2008.

[12] I.Jolliffe,PrincipalComponentAnalysis.NewYork,NY,USA:Wiley, 2005.

[13] Zhen Ye , Saurabh Prasad, Wei Li, James E. Fowler, Mingyi He, Classification Based on 3-D DWT and Decision Fusion for Hyperspectral Image Analysis, VOL. 11, NO. 1, JANUARY 2014

[14] C. Davis, Statistics and Data Analysis in Geology, 3rd ed. New York, NY, USA: Wiley, 2008, pp. 238–244.

[15] C.-C. Chang and C.-J. Lin, LIBSVM:Alibrary for support vector machines, ACM Trans. Intell. Syst. Technol., vol. 2, no. 27, pp. 1–27, 2011

[16] D. L. Donoho, "High-dimensional data analysis: The curses and blessings of dimensionality," in Proc. AMS Math Challenges Lecture, 2000, pp. 1–32.

[17] V. N. Vapnik, The Nature of Statistical Learning Theory, 2nd ed.

[18] New York, NY, USA: Springer-Verlag, 1995, pp. 138–141.

[19] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," IEEE Trans. Geosci.

[20] Remote Sens., vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[21] P. Du, K. Tan, and J. Xia, "A novel binary tree support vector machine for hyperspectral remote sensing image classification," Optics Commun., vol. 285, pp. 3054–3060, Jun. 2012.

[22] Y. Bazi and F. Melgani, "Toward an optimal SVM classification system for hyperspectral remote sensing images," IEEE Trans. Geosci. Remote Sens.,vol. 44, no. 11, pp. 3374–3385, Nov. 2006.

[23] G. C. Cawley and N. L. C. Talbot, "Over-fitting in model selection and subsequent selection bias in performance evaluation," J. Mach. Learn. Res., vol. 11, pp. 2079–2107, Jul. 2010.

[24] J. Kennedy and R. C. Eberhart, Swarm Intelligence. San Mateo, CA, USA: Morgan Kaufmann, pp. 192–206.

[25] F.Melgani and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," IEEE Trans. Inf. Technol. Biomed., vol. 12, no. 5, pp. 667–677, Sep. 2008.

[26] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning. New York, NY, USA: Springer-Verlag, 2008.

[27] H. Frohlich, O. Chapelle, and B. Scholkopf, "Feature selection for support vector machines by means of genetic algorithm," in Proc. 15th IEEE Int. Conf. Comput. Process. (Hardware/Software), Nov. 2003, pp. 142–148.

[28] S. Tang and X. Cheng, "A harmonic measuring approach based on multilayered feed forward neural network," in Proc. CSEE, 2006, vol. 26,no. 18, pp. 90–94.

[29] E. J. Mark, R. L. David, and J. H. Kastens, "Harmonic analysis of time-series AVHRR NDVI data," Photogramm. Eng. Remote Sens., vol. 67, no. 4, pp. 461–470, Apr. 2001.

[30] E. J. Mark, R. L. David, and J. H. Kastens, "Crop identification using harmonic analysis of time-series AVHRR NDVI data," Comput. Electron. Agric., vol. 37, pp. 127–139, Dec. 2002.