Automatic Digital Modulation Recognition in Presence of Noise Using SVM and PSO

M. Hadi Valipour, M. Mehdi Homayounpour and M. Amin Mehralian
Laboratory for Intelligent Multimedia Processing
Department of Computer Engineering and Information Technology
Amirkabir University of Technology (Tehran Polytechnic)
Tehran, Iran
{valipour, homayoun, mehralian}@aut.ac.ir

***Abstract*— Automatic digital modulation recognition in intelligent communication systems is one of the most important issues in software radio and cognitive radio. In this paper a new method will be presented for automatic digital modulation classification in presence of additive white Gaussian noise (AWGN). In this method a set of three different types of features is extracted to be employed in recognition process. Classification is based on support vector machine (SVM) as a powerful method for pattern recognition, and particle swarm optimization (PSO) to configure kernel parameters. Computer simulations of 16 different types of digitally modulated signals corrupted by AWGN are carried out to measure the performance of the method. Employing multiple SVMs in a hi-erarchical structure as inter-class and intra-class classifiers and also our proposed method for feature selection based on features impact on severance, presents good results in simulations. The results show that with infinite SNR, accuracy tends to 99.9%. Also this method shows eligible robustness in presence of noise as we can see in experiments conducted using low SNR data.**

Keywords- Automatic Digital Modulation Recognition (ADMR), Support Vector Machine (SVM), Particle Swarm Optimization (PSO)

I. INTRODUCTION

According to incremental expanse of intelligent receivers, automatic modulation recognition becomes a challenging topic in telecommunication systems and computer engineering. Such systems have many civil and military applications. Moreover, blind recognition of modulation type is an important problem in commercial systems, especially in software defined radio. Usually in such systems, there are some extra information for system configuration, but considering blind approaches in intelligent receivers, we can reduce information overload and increase transmission performance.

A review on previous works in this issue shows that there are two main approaches to automatic modulation recognition. The first approach uses likelihood-based methods to assign an input signal to a proper class. Another recent approach is based on feature extraction. The proposed approach in this paper is also feature-based. Below we review some related previous works.

In [1] an algorithm based on high order cumulants (HOC) and support vector machines for modulation recognition of digital communication signals is proposed. The presented method can identify six digital modulation signals: 2ASK, 4ASK, 8ASK, 4PSK, 8PSK and 16QAM digital signals, using fourth and sixth order cumulants of the signals as vectors and SVM based on binary tree as classifier. The main point of this work is a new class distance which is employed as rules of constructing binary tree, which separates the furthest class from others first.

In [2] the authors use a feature-based method introducing new intuitive features for real-time classification of digitally modulated signals without any prior knowledge of signal parameters. The incoming signal's basic modulation type is detected as FSK, PSK, ASK, QAM or GMSK and then its order is identified. The central frequency estimation method employed in this work is based on Smoothed Periodogram given by [6].

The proposed algorithm in [3] is verified using higher-order statistical moments (HOM) of continuous wavelet transform (CWT) as a features set. A multilayer feed-forward neural network is proposed as a classifier. The purpose is to discriminate among different M-ary shift keying modulation schemes and the modulation order without any priori signal information. Also in this work pre-processing and features subset selection using principal component analysis (PCA) is used to reduce the network complexity and to improve the classifier's performance.

One of the old works which have been presented in [4] proposes to use constellation shape as a robust signature for digital modulation recognition. This work represents the transmitted information by the geometry of the constellation. Received information is in turn the recovered constellation shape that is deformed by noise, channel and receiver implementation. The author demonstrates that fuzzy c-means clustering is capable of robust recovery of the unknown constellation. PSK and QAM modulation types are investigated in the experiments.

It can be perceived that previous works can be divided in two categories; some of them try to increase the accuracy of classifiers while some others want to increase the variety of modulation classes, so the accuracy of recognition will be reduced, obviously. In this paper a new approach is introduced to use a complete set of three types of features; spectral-based, statistical and wavelet-based features. The feature selection method proposed in this work is based on features impact on severance. The features impact is scored by the classifier accuracy using k-fold cross validation. A dataset of 16 different types of digitally modulated signals in 4 main categories PSK, ASK, FSK and QAM/ASKPSK are carried out to measure the performance of proposed method in experiments. In both of feature selection and classification phases, parameter configuration for SVMs has been done by PSO.

This paper is organized as follows. In section 2 signal and channel models are explained briefly. Section 3 describes features which are used in classification. In Section 4 the structure of the proposed modulation recognizer is introduced. Section 5 is about employed feature subset selection method. Simulation and experimental results are presented in Section 6. Finally, conclusion is discussed in section 7.

II. SIGNAL AND CHANNEL MODELS

In the receiver, it is assumed that the signals have been pre-processed, and reached the synchronization of the carrier frequency, phase, and timing. After down conversion, the received complex baseband signals can be expressed as [2]:

$$r(t) = c(t).e^{j2\pi f_c t} + n(t) \tag{1}$$

where f_c is the central frequency, $c(t)$ is the complex envelope and $n(t)$ is the complex additive white Gaussian noise produced by AWGN channel. The modulated signal is given by:

$$m(t) = \mathrm{Re}\{c(t).e^{j2\pi f_c t}\} \tag{2}$$

One can easily obtain the $c(t)$ for ASK, PSK and QAM by the relation

$$c(t) = \sqrt{S}\sum_{k=1}^{N}(I + jQ).p(t-kT) \tag{3}$$

where $p(t)$ is the rectangular pulse of symbol period T and S is the signal power. I and Q are In-phase and Quadrature components. For FSK the $c(t)$ can be given by:

$$c(t) = \sqrt{S}\sum_{k=1}^{N} e^{j(2\pi f_i t+\theta_i).p(t-kT)} \tag{4}$$

where f_i indicates the i^{th} frequency.

The AWGN channel is a good model for many satellite and deep space communication links. The values $n(t)$ added by the AWGN channel can be considered as a Gaussian distributed random variable with zero mean which can take any real value and is independent of the channel input. A real valued Gaussian distributed random variable X with mean value μ and variance σ^2 can be characterized by its probability function:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}.e^{(\frac{-(x-\mu)^2}{2\sigma^2})} \tag{5}$$

where in this case σ^2 depends on signal to noise ratio.

III. FEATURE EXTRACTION

As mentioned before, in this paper, the recognition is based on feature extraction from baseband signal. So there is no discussion about synchronization between data transmitters, and it is assumed that the baseband signal is available for feature extraction.

A vector of 27 features is extracted for each frame of signal (2000 digital samples or 0.25s where sampling frequency is 8000 Hz); spectral-based features, statistical features and wavelet-based features.

A. Spectral-based Features

Each feature in this category is chosen either because of its successful usage in previous works or because of mathematical analysis which indicates that it should satisfy the criteria of a good feature. The features used are described below as in [5].

Maximum value of the power spectral density of the normalised-centred instantaneous amplitude:

$$\gamma_{\max} = \frac{\max|FFT(a_{cn}(i))|^2}{N_s} \tag{6}$$

where N_s is the number of samples, $a_{cn}(i) = a_n(i) - 1$ and $a_n(i) = a(i)/m_a$, $a(i)$ is the ith instantaneous amplitude and m_a is the sample mean value.

Standard deviation of the absolute value of the centered non-linear components of the instantaneous phase:

$$\sigma_{ap} = \sqrt{\frac{1}{C}\left(\sum_{a_n(i)>a_t}\phi_{NL}^2(i)\right) - \left(\frac{1}{C}\sum_{a_n(i)>a_t}|\phi_{NL}(i)|\right)^2} \tag{7}$$

where C is the number of samples in $\{\phi_{NL}(i)\}$ for which $a_n(i) > a_t$ and a_t is the threshold value for $a(i)$ below which the estimation of the instantaneous phase is very noise sensitive, also $t = i/f_s$, where f_s is the sampling frequency (8000 Hz).

Standard deviation of the direct value of the centred non-linear component of the instantaneous phase:

$$\sigma_{dp} = \sqrt{\frac{1}{C}\left(\sum_{a_n(i)>a_t}\phi_{NL}^2(i)\right) - \left(\frac{1}{C}\sum_{a_n(i)>a_t}\phi_{NL}(i)\right)^2} \tag{8}$$

Standard deviation of the absolute value of the normalised-centred instantaneous amplitude:

$$\sigma_{aa} = \sqrt{\frac{1}{N_s}\left(\sum_{i=1}^{N_s}a_{cn}^2(i)\right) - \left(\frac{1}{N_s}\sum_{i=1}^{N_s}|a_{cn}(i)|\right)^2} \tag{9}$$

Standard deviation of the absolute value of the normalised-centred instantaneous frequency:

$$\sigma_{df} = \sqrt{\frac{1}{C}\left(\sum_{a_n(i)>a_t}f_N^2(i)\right) - \left(\frac{1}{C}\sum_{a_n(i)>a_t}|f_N(i)|\right)^2} \tag{10}$$

Note that the availability of instantaneous amplitude, phase and frequency depends on the availability of carrier frequency f_c. In this paper, we assume that f_c is known a priori. If a good linear approximation of non-linear component of instantaneous phase $\phi(t)$ is available, instantaneous frequency can be derived as below:

$$f(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} \tag{11}$$

B. Statistical Features

Some of modulation schemes often contain information in both amplitude and phase spectra. In telecommunications, these modulations are regarded as complex signal and often are illustrated as in-phase and quadrature components. In this work, a different scheme by using the cumulants of the real and imaginary parts of the analytic signal is used taking inspiration from [5].

Let X_i be a signal vector, $\{x_i^1, x_i^2, ..., x_i^N\}$ and $\langle \; \rangle$ denote the statistical expectation. The second, third and fourth-order cumulants at zero lag will be computed as:

$$C_{X_1,X_2} = \langle X_1, X_2 \rangle = \frac{1}{N}\sum_{n=1}^{N} x_1^n x_2^n \tag{12}$$

$$C_{X_1,X_2,X_3} = \langle X_1, X_2, X_3 \rangle = \frac{1}{N}\sum_{n=1}^{N} x_1^n x_2^n x_3^n \tag{13}$$

$$\begin{aligned} C_{X_1,X_2,X_3,X_4} &= \langle X_1, X_2, X_3, X_4 \rangle \\ &- \langle X_1, X_2 \rangle \langle X_3, X_4 \rangle - \langle X_1, X_3 \rangle \langle X_2, X_4 \rangle \\ &- \langle X_1, X_4 \rangle \langle X_2, X_3 \rangle \\ &= \frac{1}{N}\sum_{n=1}^{N} x_1^n x_2^n x_3^n x_4^n - C_{X_1,X_2} C_{X_3,X_4} \\ &- C_{X_1,X_3} C_{X_2,X_4} - C_{X_1,X_4} C_{X_2,X_3} \end{aligned} \tag{14}$$

Also we need to describe the complex envelope of the sampled signal which is defined by

$$H_y = [y(t) + j\hat{y}(t)]\exp(-i2\pi f_c t) \tag{15}$$

where $\hat{y}(t)$ is the Hilbert transform of $y(t)$.

The proposed statistical feature set includes $\{C_{R,R}, C_{R,I}, C_{I,I}, C_{R,R,R}, C_{R,R,I}, C_{R,I,I}, C_{I,I,I}, C_{R,R,R,R}, C_{R,R,R,I}, C_{R,R,I,I}, C_{R,I,I,I}, C_{I,I,I,I}\}$, where R is defined to be the real part of H_y, and I is the imaginary part of H_y.

C. Wavelet-based Features

Another feature type which also employed in [3] is a set of higher-order statistical moments of continuous wavelet transform. The continuous wavelet transform of a received signal x(t) is defined as [7]

$$CWT(a,\tau) = \int_{-\infty}^{\infty} x(t)\psi_{a,\tau}^{*}(t)dt \tag{16}$$

where $a > 0$ is the scale variable, $\tau \in R$ is the translation variable, and * denotes complex conjugate. This defines the so-called CWT, where CWT(a,τ) defines the wavelet transform coefficients. The Haar wavelet is chosen as the mother wavelet.

Two sets of features are extracted here and each of them has 6 elements. The first set includes 1^{st} to 6^{th} order statistical moments of CWT and the second set includes 2^{nd} to 6^{th} order statistical moments of CWT which a median filter has been applied to cut off the peaks in the corresponding wavelet transforms before. These 2 sets will be shown as $HOM_{order}(CWT(a,\tau))$ and $HOM_{order}(CWT(a,\tau)_{filtered})$. The absolute value of continuous Haar wavelet transform of 8FSK and 8PSK signals are shown in Fig. 1.

IV. MODULATION CLASSIFICATION

As it was said before, in this work an SVM classifier is used in a hierarchical structure and PSO is used to configure parameters. SVM classifier and PSO are briefly described in the following sections.

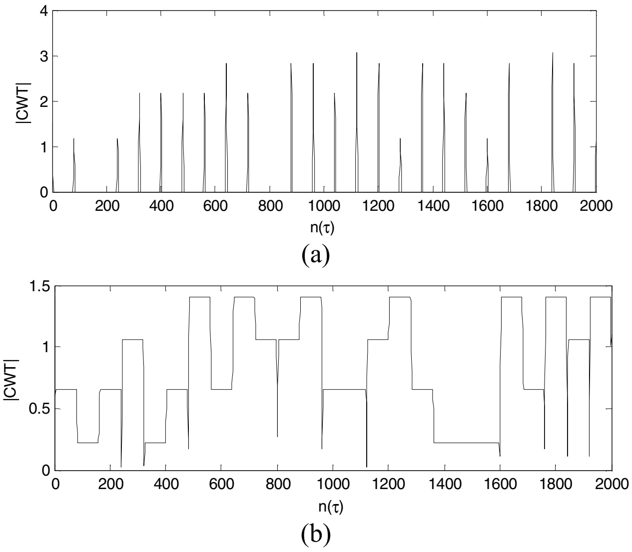

Figure 1. The absolute value of continuous Haar wavelet transform of (a) 8PSK and (b) 8FSK

A. Support Vector Machine

SVM is an empirical modeling algorithm that can be applied in classification problems. The first objective of the Support Vector Classification (SVC) is the maximization of the margin between each two nearest data points belonging to two separate classes. The second objective is to constrain that all data points belong to the right class [8] [11].

In the proposed method, SVM is used in a hierarchical structure as can be seen in Fig. 2. In this structure at the first layer, the incoming signal's modulation type is detected as FSK, PSK, ASK or QAM/ASKPSK and then in the second layer its order is identified by corresponding classifier. Therefore there are 1 inter-class classifier and 4 intra-class classifiers. We used one against all method in multi-class classification.

One of the most important problems in these classifiers is their sensitivity to kernel parameters. In this paper to find the proper parameters for a Radial Basis Function (RBF) kernel, PSO can be employed as a popular optimizer in similar issues.

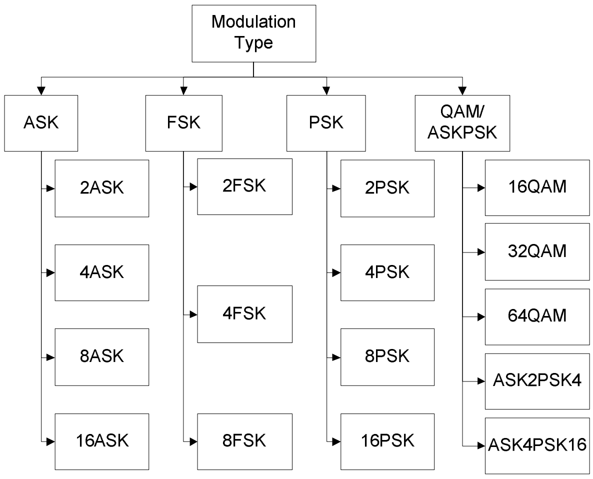

Figure 2. Hierarchical structure of classifier

B. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a technique used to explore the search space of a given problem to find the settings or parameters required to maximize a particular objective. This technique, first described by James Kennedy and Russell C. Eberhart in 1995, originates from two separate concepts: the idea of swarm intelligence based on the observation of swarming habits by certain kinds of animals (such as birds and fishes) and the field of evolutionary computation. The PSO algorithm consists of just three steps, which are repeated until some stopping condition is met: First, evaluation of the fitness of each particle, second, updating individual and global best fitness and positions and third, updating velocity and position of each particle. [9].

Each particle position will be updated as below:

$$X_{new} = X_{old} + V_{new} \tag{17}$$

V is new velocity which can be derived as

$$\begin{aligned} V_{new} &= V_{old} + c_1 \times rand_1(pBest - X) + \\ &c_2 \times rand_2(gBest - X) \end{aligned} \tag{18}$$

where the *pBest* is the individual best candidate solution for particle, and *gBest* is the swarms global best candidate solution. c_1 and c_2 are learning coefficients. Random coefficients are used to keep the diversity in solution space. Here the value of each particle is based on 5-fold cross validation accuracy of classifier.

V. FEATURE SELECTION

As it was introduced before, the extracted dataset includes 27 different features, so the complexity and runtime will be increased in computations and also the generalization will be reduced. Therefore a feature subset selection process is needed. In current research this process is based on classifier accuracy and it is assumed that all the features are independent. The impact of each feature is scored by the accuracy of the classifier which employs just this feature. This process has been done 5 times for each classifier (inter-class and intra-class classifiers) and the mean value of results is considered. In the next phase, features will be ranked by their scores and the best ranked features will be selected. Table 1 shows selected features for each classifier.

Fig. 3 shows a schema of feature extraction, feature selection, classifier training and parameter configuration, test phase in the proposed approach.

VI. SIMULATIONS AND RESULTS

In this section the performance of the proposed approach is measured by conducting some experiments. Matlab Communication Toolbox is used to produce various types of modulated signals with different SNRs. Also the LibSVM [10] toolbox based on Matlab environment is used for multiclass training and testing of support vector machines.

TABLE I. SELECTED FEATURES FOR EACH CLASSIFIER

intra-class classifier	FSK classifier	PSK classifier
σ_{dp}	σ_{dp}	$HOM_2(CWT(a,\tau))$
σ_{ap}	σ_{ap}	$HOM_3(CWT(a,\tau))$
σ_{af}	σ_{af}	$HOM_4(CWT(a,\tau))$
σ_{aa}	σ_{aa}	$HOM_5(CWT(a,\tau))$
$HOM_2(CWT(a,\tau))$	C_{II}	$HOM_6(CWT(a,\tau))$
$HOM_3(CWT(a,\tau))$	$HOM_3(CWT(a,\tau))$	$HOM_2(CWT(a,\tau)_{filtered})$
$HOM_2(CWT(a,\tau)_{filtered})$	$HOM_3(CWT(a,\tau)_{filtered})$	$HOM_3(CWT(a,\tau)_{filtered})$

ASK classifier	QAM/ASKPSK classifier
$HOM_2(CWT(a,\tau))$	$HOM_2(CWT(a,\tau))$
$HOM_2(CWT(a,\tau)_{filtered})$	$HOM_2(CWT(a,\tau)_{filtered})$

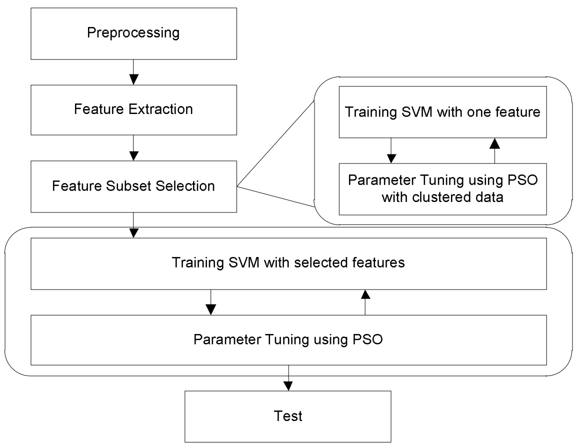

Figure 3. Schema of the proposed automatic modulation recognition approach

In all of experiments it is assumed that the center frequency is available (1800 Hz). As it said in the introduction section, different parameters are considered to have a diverse dataset. Different symbol rates including 50, 75, 96, 100, 125, 200, 300, 600, 800, 1280, 1500 and 1800 baud are employed. Also in PSK and QAM there is a roll off parameter which is configured to be 0.0, 0.1, 0.2 and 0.35 in signal generation.

Modulated signals are corrupted by AWGN at four different SNRs: 6_{dB}, 9_{dB}, 16_{dB} and infinite.

In our experiment, the proposed classifier is trained using data at infinite SNR and then tested with corrupted data. The aim of this experiment is to evaluate the robustness and accuracy of the introduced method in AWGN channels. In the second phase, training is also carried out using more amounts of corrupted data by AWGN. Generated dataset contains 2000 feature vectors for each modulation, which 70% has been used for training and 30% for testing. Also 10% of whole dataset selected randomly, has been used for parameter tuning by PSO.

Fig. 4 shows the classification accuracy at different SNRs for each classifier and Fig. 5 depicts the classification accuracy at different SNRs for proposed method compared to other recent approaches [3] [5] for modulation recognition .

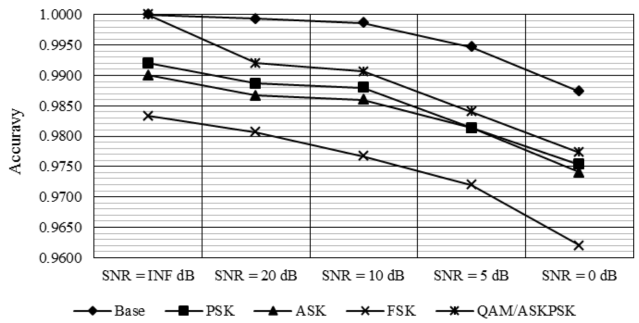

Figure 4. Classification accuracy at different SNRs for each classifier

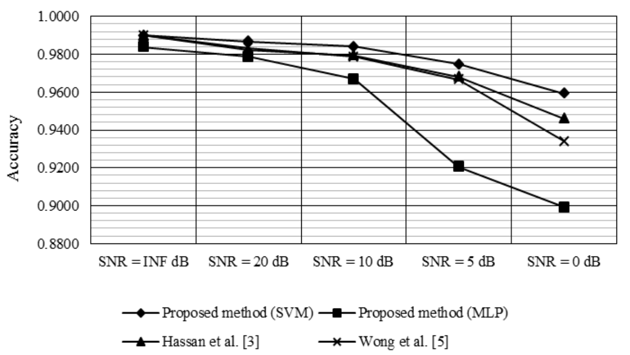

Figure 5. Classification accuracy at different SNRs for proposed method comparing other recent approaches for modulation recognition

The simulation results that the accuracy rate of this method is at least 96% at 0dB SNR using enough and varied training data, while the accuracy is 94.5% and 93.5% for recently proposed Hassan et al. [3] and Wong et al. [5] methods respectively. Also the automatic recognition method exhibits a satisfying performance at low SNRs even as low as 0_{dB}. Feature selection for inter-class classifier leads to very low miss classification rate.

VII. CONCLUSION

In this paper a new method based on Support Vector Machines and Particle Swarm Optimization was introduced for automatic digital modulation recognition. In the proposed method, the classifier was used in a hierarchical structure. This structure includes 1 inter-class classifier and 4 intra-class classifiers. At the first layer the incoming signal's modulation type was detected as one of four main modulation classes and then in the second layer its order was identified by corresponding classifier. Feature subset selection was carried out for each classifier. 16 types of modulation schemes were used in the experiments as well as different parameters employed in signal generation. It can be perceived from experiment results that the proposed method is in a good trade of between the recognition accuracy and the number of modulation types. Also the experiment depicts the robustness of the proposed method in presence of AWGN noise.

REFERENCES

[1] L.-X. Wang, Y.-J. Ren, and R.-H. Zhang, "Algorithm of digital modulation recognition based on support vector machines," in 2009 International Conference on Machine Learning and Cybernetics, vol. 2, 2009, pp. 980-983.

[2] S. M. Baarrij, F. Nasir, and S. Masood, "A Robust Hierarchical Digital Modulation Classification Technique: using Linear Approximations," in 2006 IEEE International Symposium on Signal Processing and Information Technology, pp. 545-550, 2006,.

[3] K. Hassan, I. Dayoub, W. Hamouda, and M. Berbineau, "Automatic Modulation Recognition Using Wavelet Transform and Neural Networks in Wireless Systems," EURASIP Journal on Advances in Signal Processing, pp 1-13, 2010.

[4] B. G. Mobasseri, "Digital modulation classification using constellation shape," Signal Processing, vol. 80, no. 2, pp. 251-277, 2000.

[5] M. Wong and A. Nandi, "Automatic digital modulation recognition using artificial neural network and genetic algorithm," Signal Processing, vol. 84, pp. 351-365, 2004.

[6] P. Bloomfield, Fourier analysis of time series: an introduction, 2nd ed., New York: John Wiley & Sons, Inc., USA, 2000.

[7] Charles K. Chui, An Introduction to Wavelets. Academic Press Prof., Inc., San Diego, CA, USA, 1992.

[8] Christopher J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data Min. Knowl. Discov. Vol. 2, No. 2, pp 121-167, June 1998.

[9] Riccardo Poli, James Kennedy and Tim Blackwell, "Particle swarm optimization, An overview," Swarm Intelligence, Vol. 1, No. 1, pp. 33-57, June 2007.

[10] Chang CC, Lin CJ, "LIBSVM: a library for Support Vector Machines," Tech. rep., Department of Computer Science, National Taiwan University, 2003.

[11] Hussam Mustafa and Miloš Doroslovaki, "Digital Modulation Recognition Using Support Vector Machine Classifier," in Proc. of Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, Vol.2, pp 2238 - 2242, 2004.