Principal Component Analysis of Cyclic Spectrum Features in Automatic Modulation Recognition

Fangming He, Yafeng Yin, Lei Zhou, Xingzhong Xu, Hong Man
Department of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, NJ, 07030
Email: {fhe,yyin1,lzhou,xxu7,hman}@stevens.edu

***Abstract*—Automatic modulation recognition (AMR) of communication signals is a critical and challenging task in cognitive radio systems. In this work, classifications of four digital modulation types, including BPSK, QPSK, GMSK and 2FSK, are investigated. From the received radio signal, a set of cyclic spectrum features are first calculated, and a principal component analysis (PCA) is applied to extract the most discriminant feature vector for classification. A novel max-multiple layer perceptron (MaxMLP) neural network is introduced for classification of modulation feature vectors through supervised learning. In the experiments, real radio signals with different modulation types were generated from an Agilent vector signal generator, and sampled by an Agilent digital signal analyzer. The proposed AMR method is tested at various channel SNR levels. Experimental results indicate that the performance of this method is highly competitive, and the computational cost is relatively low.**

I. Introduction

Cognitive radio (CR), built upon the software-defined radio (SDR) technology, is perceived as an intelligent wireless communication system that is able to sense the environment, make decisions based on the observations, and learn from experience to improve future decisions [1]. The most important characteristic of the cognitive radio is that it can recognize and adapt to different requirements of communication systems and various environments. To accommodate to different communication systems, particularly the military and emergency response systems, the capability of automatic modulation recognition (AMR) of radio signals is essential. Without any prior information at the receiver, such as pilot, timing information, carrier frequency, phase offsets and so on, blind identification of the modulation type is a challenging task.

Modulation classification can be categorized into two major groups, i.e. decision-based (DB) methods and feature-based (FB) methods. The former approach is based on the likelihood functions of received signals, and pre-defined decision thresholds are determined by comparing the likelihood ratios. However, the dynamic nature of the radio environment for most communication systems makes the likelihood ratio test ineffective. In the feature-based approach, different features are extracted from signals and a suitable classifier is applied to partition the feature space and identify different modulation types. Many spectrum features have been extensively studied for modulation classification, including instant waveform and spectrum [2], high-order statistics information [3], cyclostationarity [4] and wavelet transform [5]. Extracting a proper set of features for classification also has many practical issues in real applications. For example, without prior knowledge, the instantaneous phase or frequency cannot be correctly estimated. Therefore, recent papers focus on exploiting high-order statistics and cyclostationarity features [6], [7]. The high-dimensional cyclostationarity features effectively capture the distinctive cyclic characteristics of many modulation types. However, the high feature dimension prevents it from efficient implementation. Authors in [6], [8], [9] took the maximum value along the spectral domain as a simplified cyclostationarity feature set for modulation classification, which was called Cyclic Domain Profile (CDP). It greatly reduces the feature dimensions without sacrificing too much performance of classification. Besides feature selection, classifier design also plays a essential role in modulation classification. Artificial intelligence (AI) algorithms, such as artificial neural networks (ANN) and some other machine learning methods, have been widely employed and promising recognition results have been reported [10].

In this work, we introduce a new feature extraction scheme, which applies principal component analysis (PCA) on the cyclostationarity features and obtains the first one or more principal components to form the principle cyclic spectrum (PCS) feature vector. The PCS feature vectors have much lower dimension than the original cyclostationarity features, and yet they preserve the major distinctive characteristics of different modulation types. In addition, we develop a new MaxMLP neural network architecture, in which four multilayer perceptron (MLP) classifiers [11] are trained for four parts of modulation data respectively. Based on the majority voting of the four classifiers, the final output is jointly decided. In order to study the efficiency of our proposed AMR algorithm, a test bench consisting of an Agilent's vector signal generator (VSG) and a digital signal analyzer (DSA) is established. All training and test data samples are captured from real radio signals.

The rest of the paper is organized as the follows. The extraction of PCS feature vectors is introduced in section II. The feedforward MaxMLP neural network is introduced in section III. The experiment set up and performance evaluation are presented in section IV. Finally, conclusions drawn from this work are provided in section V.

II. Principal component analysis on cyclostationary spectral

Extracting highly distinctive features is essential for modulation classification. According to previous work in [4], the modulated signal can be considered as a cyclostationary process and the spectral coherence function can be calculated as the feature for AMR. To make the computation feasible in the classification phase, authors in [6] introduced the CDP features and demonstrated good performance in AMR. However, the dimension reduction from the original cyclostationarity features to the CDP features is not optimized in any sense, and causes major loss of discriminant information, especially in a highly noisy environment. Therefore we propose to apply principal component analysis on the spectral coherence function for dimension reduction, which produces a new principle cyclic spectrum (PCS) feature vector. PCS has a lower dimensional representation of the spectral coherence, and it preserves the most important spectral information for modulation classification in the sense of maximum energy compaction.

A. Cyclostationary spectral function

Cyclic spectral coherence, which deals with the second order transformation of a function and its spectral representation, has been used as an efficient feature to determine the modulation type of unknown signals. It is well argued that the modulated signal may be described as a cyclostationary process, which varies periodically. This periodicity, resulting from periodic sampling, scanning, modulating, multiplexing and coding, can be exploited to determine the modulation format of the unknown signals. For a received modulated signal $x(t)$, we may describe the autocorrelation function as

$$R_x(t,\tau) = E[x(t+\tau/2)x^*(t-\tau/2)] = \sum_{\alpha} R_x^{\alpha}(\tau)e^{j2\pi\alpha t}. \tag{1}$$

Here, E is the expectation operator; α is the set of Fourier components, which can be expressed as

$$R_x^{\alpha}(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} R_x(t,\tau)e^{-j2\pi\alpha t}dt. \tag{2}$$

The autocorrelation function is then periodic in t with period T_0

$$R_x^{\alpha}(\tau) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} R_x(t,\tau)e^{-j2\pi\alpha t}dt. \tag{3}$$

According to the Wiener-Khintchine theorem, the spectral coherence function (SCF) is the Fourier transform of the autocorrelation function, and it is given by

$$S_x^{\alpha}(f) = \int_{-\infty}^{\infty} R_x^{\alpha}(\tau)e^{-j2\pi f\tau}d\tau. \tag{4}$$

The SCF is cross-correlated between two frequency components separated by $f-\alpha/2$ and $f+\alpha/2$. If $x(t)$ does not include spectral components at $f=\pm\alpha/2$, then the SCF becomes the covariance of the two spectral components, which means we may normalize it by the geometric mean. Therefore, the normalized SCF may be given by

$$C_x^{\alpha}(f) = \frac{S_x^{\alpha}(f)}{[S_x^0(f+\alpha/2)S_x^0(f-\alpha/2)]^{\frac{1}{2}}} \tag{5}$$

The range of the normalized SCF falls in [0,1]. The cyclostationarity features of the BPSK, QPSK, GMSK and 2FSK are illustrated in Fig. 1.

The cyclic spectral coherence is widely used to classify modulated signals. However, the high dimentionality of the SCF representation makes it difficult to be directly applied in the signal identification algorithms. Some researchers [6], [8], [9] proposed the idea of projecting the SCF of the signal to cyclic domain profile and picking the peak values to determine the modulation format. This method is not reliable when the SNR is low, because some important information is lost. In order to decrease the classification errors in low SNR scenarios and to reduce computational complexity, we apply the PCA method to these spectral coherence features. In this way, we can preserve most discriminant information and greatly decrease the required feature dimensions in classification.

B. Principal component analysis on signal

PCA is an orthogonal linear transform which projects the data to a new coordinate space [12]. It is a popular dimension reduction transform that retains most of information of the original data. It is a theoretically optimum transform for a given data set in terms of least mean square errors.

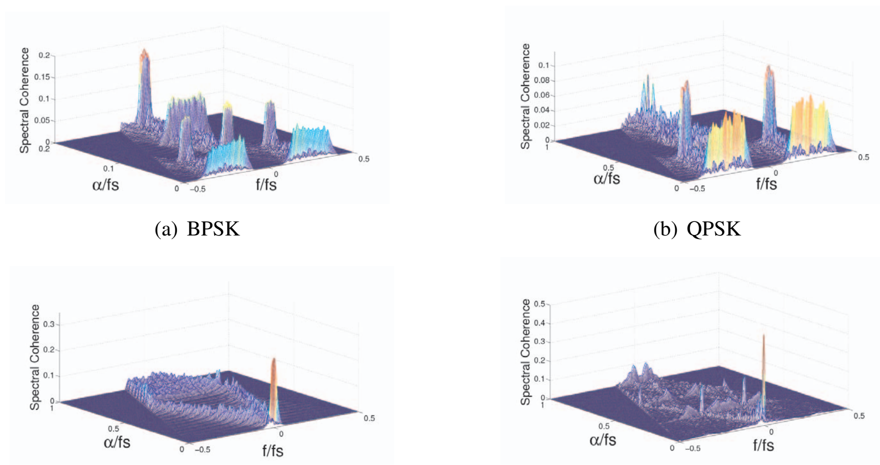
Assume we have N independent normalized SCFs at cyclic frequency α as $C_x = [c_{f_1},\ldots,c_{f_S}]$, and for each separated frequency f_i, there are $c_{f_i} = [c_{f_i}^{\alpha_1},\ldots,c_{f_i}^{\alpha_N}], i=1,\ldots,S$. The features have latent patterns but are hard to describe in the high dimension. PCA maps a given set of data points onto the principal components ordered by the amount of data variance that they capture. The first m principal components $\mathbf{V}=\{v_i\}(i=1,\ldots,m)$ are

$$v_i = argmax_{\|x\|=1} \parallel (\mathbf{C} - \sum_{j=1}^{i-1}\mathbf{C}v_jv_j^T)x \parallel, \tag{6}$$

which is the i_{th} eigenvector of the estimated covariance matrix $\mathbf{C}$ of C_x, with $\mathbf{C}=E[(C_x-E[C_x])(C_x-E[C_x])^T]$. Therefore, PCA can be viewed as a SVD of the covariance matrix. The key idea behind it is that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on.

PCA will retain k orthogonal vectors as the principal components, which capture the highest variance in the original matrix ("normal space"). By extracting these k vectors, the remaining $m-k$ vectors can be considered as statistical anomalies or noise ("residual space").

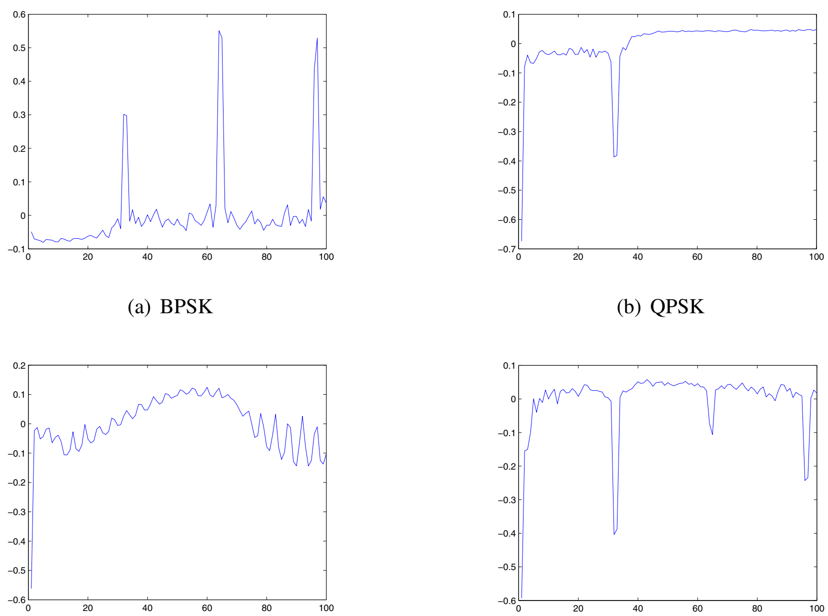
The first principal component of the SCF is illustrated in Fig.2.



(a) BPSK (b) QPSK

(c) GMSK (d) 2FSK

Fig. 1. Spectral coherence of the modulated signals with SNR 20dB



(a) BPSK (b) QPSK

(c) GMSK (d) 2FSK

Fig. 2. The first order of the principal component of spectral coherence with SNR 20dB

III. AMR BASED ON PCS

Artificial Neural Network (ANN) is one of the most popularly used tools in machine learning applications. It has been extensively studied in various cognitive radio applications, such as spectrum sensing, modulation classification and demodulation. A typical ANN is comprised of sets of connecting neurons with different network structures. In the training stage, a series of neurons' weights will be adjusted to capture the underlying relationship between the input vector and the desired output. In general, ANN can be used to approximate any spatially finite functions by different network structures. According to different connecting structures, ANN can be categorized to three groups: feedforward, feedback and self-organizing neural networks. In this paper we focus on the feedforward (with error back propagation) neural network for their simplicity in the modulation classification.

A. *Multiple Layer Perception Neural Network*

Multiple Layer Perception (MLP) neural network is a widely used feedforward neural network, which is comprised of layers of neurons. In an MLP, each neuron is connected by the linear combination of neurons' outputs from the previous layer. The weight is initialized with a random value and will be updated by different optimization methods (such as genetic algorithm, scale conjugate gradient, etc) during the training process. A trained MLP then can be applied to classify different input features vectors through the feedforward procedure.

B. MaxMLP Network

Inspired by MAXNET [6], we propose a novel MaxMLP neural network structure for modulation classification in this paper. For each modulation type in the MAXNET neural network, a corresponding binary classifier will be trained for this modulation versus all the other modulation types. Thereafter a number of binary classifiers will be built after the training phase. In the testing phase, the maximum output of these binary classifiers will be selected as the final result. In the proposed MaxMLP neural network, a bagging mechanism is employed for classification as each neural network can be treated as a weak classifier. MaxMLP is comprised of multiple identical MLPs and each MLP is a multiple-class classifier rather than a binary classifier. The MaxMLP structure is shown in Fig.3. During the learning stage, each MLP is trained with different data which is randomly sampled from the whole data set. The testing data will be feeded forward to all the trained MLPs in parallel, and the final classification result will be taken through a majority voting mechanism of all the outputs of the classifiers. In our MaxMLP, each MLP classifier is made

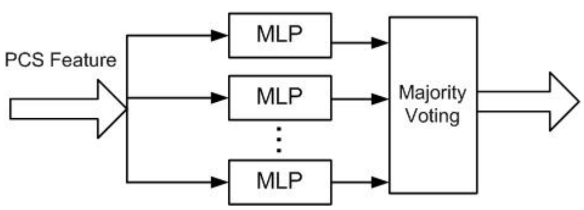


Fig. 3. MaxMLP structure

up of 5 hidden nodes with a learning rate $\mu = 0.01$. The limit of the epoches is 100 for every sub-classifier.

IV. Experiments and Performance Evaluation

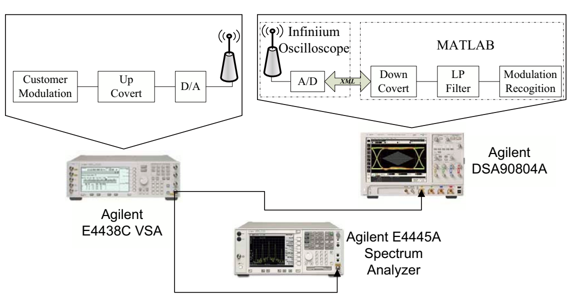


Fig. 4. The system architecture of the modulation recognition

In this project, an Agilent E4438C vector signal generator (VSG) is used as the signal source to generate required modulation types. An Agilent E4445A spectrum analyzer is used to verify the generated signals. For the receiver, we utilize an Agilent 90804A Digital Oscilloscope Analyzer (DSA) to implement high rate sampling. As shown in Fig.4, Agilent's 90804A DSA integrates the Infiniium Oscilloscope and MATLAB software. The Infiniium has the capability to observe fast varying signals and save long sequences of signal values. The DSA provides 8 GHz bandwidth on each of the four channels, with a maximum 40G/seconds sampling rate on those channels. The standard memory (10 Mpoints) on four channels allows a storage of up to 1 Gigapoints on four channels. For our AMR project, the DSA allows direct acquisition of those modulated signals without the complexity of designing a very high speed A/D converter.

In order to quickly evaluate our algorithms, we utilize XML to create a direct gateway between the DSA and MATLAB. Through XML programming, we describe the user interface and provide the data stream pipeline between them. The XML file contains tags and field values, from which we add functions of our algorithm as a math operator entry in the dialog box of the oscilloscope in the DSA. The captured waveform data is passed to a MATLAB script that we use to calculate the features and execute the AMR algorithms. The calculated features and results of the AMR are passed back to the oscilloscope to be displayed as a function in the waveform viewing area. Fig.4 illustrates the architecture of our system. When the DSA acquires the live waveform, the data is processed by the way of software defined radio components, including an RF signal downconverter, low pass filters and the AMR algorithm implemented in MATLAB.

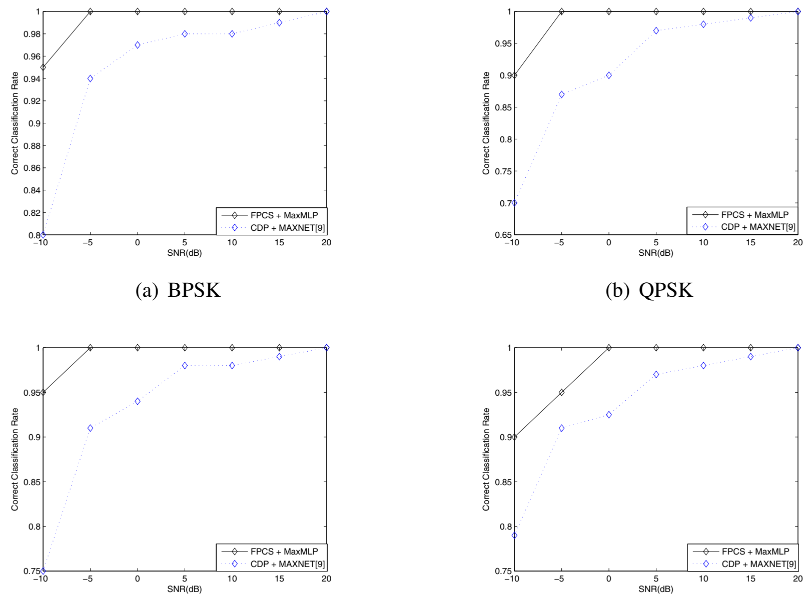
In this platform, all the modulated signals share the same symbol rate (10K/second), sampling rate (40K/second) and carrier frequency (1MHz). The bandwidth of the low pass filter is 70kHz, which is higher than the bandwidth of the received modulated signals.

Experiments were run with the same number of BPSK, QPSK, GMSK and 2FSK modulated signals. In all experiments, we inspect performance at different SNRs from -10dB to 20dB assuming no prior knowledge of the received signal's carrier frequency, timing information and symbol rate. The Max-MLP neural networks are trained with cross validation. Four different scenarios were examined in our experiments. In the first three experiments, 1200 PCSs are used to train our system. In the forth experiment, the number of training PCS samples is examined at different SNRs.

First, we analyze the performance of our proposed scheme in comparison with a MAXNET based AMC algorithm using CDP features, as introduced in [9].

In Fig.5, the black solid lines illustrate the performance of our proposed MaxMLP neural networks with the first principal component of spectral coherence feature (FPCS); the blue dot lines indicate the performance of MAXNET neural networks with CDP feature [9]. From Fig. 5, we see that our proposed AMR algorithm always achieves better performance than that of [9]. For our proposed AMR algorithms, the correct classification rate of any one of the modulation schemes is over 90% even when the SNR is as low as -10dB, while the successful classification rate of the MAXNET with CDP features is about 75 % when the SNR is -10dB.

In the second experiment, we examine the efficiency of our proposed MaxMLP neural networks in comparison with the



(a) BPSK (b) QPSK

(c) GMSK (d) 2FSK

Fig. 5. The MaxMLP with PCA feature V.S. MAXNET with CDP feature

MAXNET. From Fig.6, it can be observed that the MaxMLP neural networks are superior to the MAXNET neural networks when both are using the same PCS features. Comparing Fig.5 with Fig.6, we may find that the performance of the MAXNET with the PCS features is better than that of the same MAXNET with the CDP features.

In the third experiment we investigate the effect of different number of principal components on the modulation classification performance, using the proposed MaxMLP neural network with PCS features.

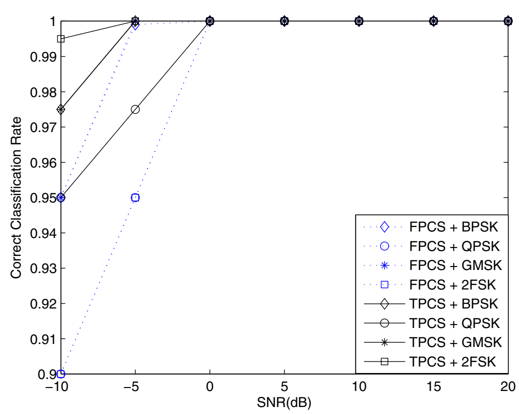


Fig. 7. The classification rate under the first one component V.S. that of the first two principal components

From Fig.7, the classification rates with the first principal component of cyclic spectral is not as high as that with the first two principal components of cyclic spectral (TPCS). However, the computational complexity with two principal components of SCF is doubled in comparison with the one using only the first principal component. There is a tradeoff between the computational complexity and the required correct classification rate for the specific SNRs. With an increase in the number of principle components, the performance of the classification will increase. When the number of the components reach some specific value, the improvement will be limited. For example, the correct classification rate will arrive to 95% even when the SNR is only -10dB.

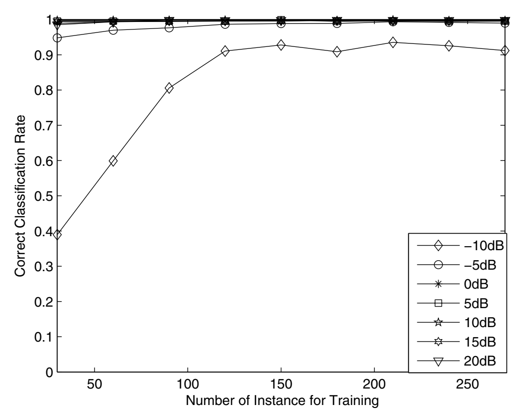
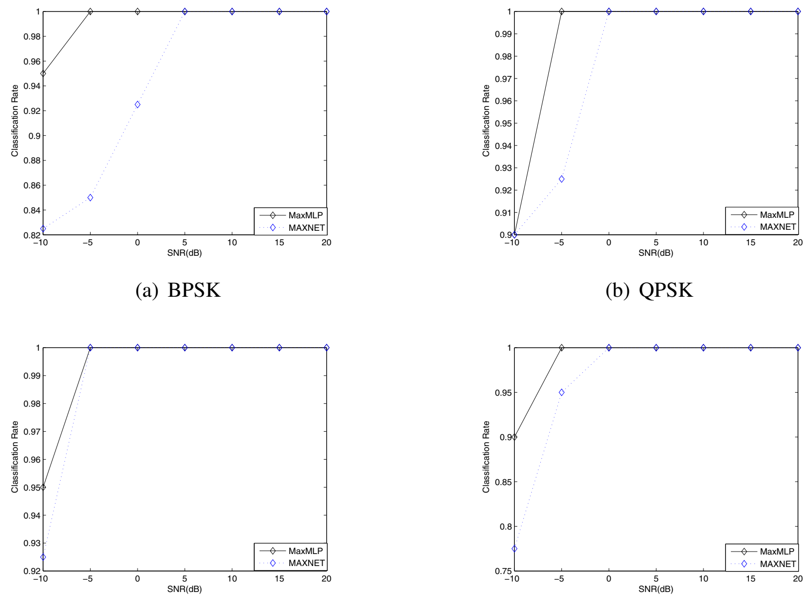


Fig. 8. Probability of correct classification versus number of training instances

Finally, the reliability of our AMR algorithm is evaluated in the forth experiment. The required number of instances for training the MaxMPL neural networks is explored. Without enough training instances, the neural networks cannot be successfully built. On the other hand, too many instances may also result in over-fitting, which may yield some unnecessary



(a) BPSK (b) QPSK

(c) GMSK (d) 2FSK

Fig. 6. The MaxMLP V.S. MAXNET with our proposed PCS feature

testing error. Fig.8 illustrates that the required instances for training a stable MaxMLP neural network for the Low SNR is higher than that for the High SNRs. For example, 120 instances with -10dB are required for our MaxMLP neural networks to reach a stable classification rate, while only 30 instances with 0 dB are required to get to a stable classification rate.

V. Conclusion

In this paper, we proposed a method for classification of communication modulation types based on a new MaxMLP neural network. The principal component features of the cyclic spectral of the modulated signal is also introduced. By extracting the principal component features of the cyclic spectral, we may reduce the computational complexity without sacrificing the performance of modulation classification. The higher the number of the principal components is, the higher the performance of the modulation classificaiton becomes.

Compared with the previous works based on the cyclic domain profile of the cyclic spectral, our AMR method without any priori knowledge can achieve a higher classification rate as well as decrease the computational complexity. We run many experiments to validate the performance of our AMR method. The effects of the number of the components and the number of training instances on the classification rate are also examined.

Acknowledgments

The authors would like to thank Prof. Joseph Mitola III and Prof. Yu-dong Yao for posing the challenge problems and instructions on this project. Fruitful suggestions and comments with Dr. Mary R. Schurgot and Prof. Brunce. McNair are also gratefully acknowledged.

References

[1] J. Mitola and M. Q. Gerald, “Cognitive radio: making software radios more personal,” *Personal Communications, IEEE*, vol. 6, no. 4, pp. 13–18, August 1999.

[2] W. Zong, E. M.-K. Lai, and C. Quek, “Digital modulation classification using fuzzy neural networks,” in *Chance Discoveries in Real World Decision Making*, 2006, pp. 101–116.

[3] M. L. D. Wong and A. K. Nandi, “Automatic digital modulation recognition using artificial neural network and genetic algorithm,” *Signal Process.*, vol. 84, no. 2, pp. 351–365, 2004.

[4] E. Like, vasu D. Chakravarthy, P. Ratazzi, and Z. Wu, “Signal classification in fading channels using cyclic spectral analysis,” *EURASIP Journal on Wireless Communications and Networking*, July 2009.

[5] P. Prakasam and M. Madheswaran, “Digital modulation identification model using wavelet transform and statistical parameters,” *J. Comp. Sys., Netw., and Comm.*, vol. 2008, pp. 1–8, 2008.

[6] A. Fehske, J. Gaeddert, and J. Reed, “A new approach to signal classification using spectral correlation and neural networks,” in *Dynamic Spectrum Access Networks, 2005. DySPAN 2005*, 2005, pp. 144 – 150.

[7] O. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, “Survey of automatic modulation classification techniques: classical approaches and new trends,” *Communications, IET*, vol. 1, no. 2, pp. 137–156, April 2007.

[8] W. C. Headley, J. D. Reed, and C. R. Silva, “Distributed cyclic spectrum feature-based modulation classification,” in *Wireless Communications and Networking Conference(WCNC)*, April 2008, pp. 1200–1204.

[9] B. Ramkumar, “Automatic modulation classification for cognitive radios using cyclic feature detection,” *IEEE Circuits and Systems*, vol. 2, no. 2, pp. 27–45, Second Quarter 2009.

[10] A. He, K. K. Bae, T. R. Newman, J. Gaeddert, K. Kim, R. Menon, L. Morales, J. Neel, Y. Zhao, J. H. Reed, and W. H. Tranter, “A survey of artificial intelligence for cognitive radios,” *IEEE Transactions on Vehicular Technology*, 2010.

[11] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Mathematics of Control, Signals, and Systems*, vol. 2, pp. 303–314, 1989.

[12] J. Shlens, “A tutorial on principal component analysis,” in *Systems Neurobiology Laboratory, Salk Institute for Biological Studies*, 2005.