Dominant Eigen Vector Based Feature Selection Using Singular Value Decomposition in Automatic Modulation Classification

Afan Ali and Fan Yangyu

Abstract—**Feature based methods for Automatic Modulation Classification (AMC) have been a widely studied topic. This paper strives to design a methodology for feature selection using t-test statistics and Singular Value Decomposition (SVD) based dominant Eigen vectors. It then investigates the performance of K Nearest Neighbor (KNN) and Multiclass OnevsAll (OVA) Support Vector Machine (SVM) using the selected feature vector. Features are generated using Ambiguity Function (AF) of the modulated signals. Extensive cross validation is done to check the feature selection algorithm. Results show that Multiclass SVM classifier gives slightly better performance than KNN classifier.**

Index Terms—**Automatic modulation classification, feature extraction, singular value decomposition, classifier performance, support vector machine, k nearest neighbor, eigen vectors.**

I. Introduction

Communication intelligence and Electronic Support (ES) system requires a very important signal processing sub module known as Automatic modulation classification (AMC) [1]. Lately, effective classification of modulation is not only needed in civilian applications like parameter estimation and signal recognition [2], but also aids many military application esp. in signal interception, jamming and signal identification [3].

AMC usually takes place between signal detection and demodulation [1]. Literature narrows down two main sub types of AMC algorithms: likelihood-based (LB) [4]-[7] and feature-based (FB) [8]-[13] methods. LB method computes a likelihood function of the received signal and then makes a decision by comparing it to a threshold. On the other hand, FB method computes a number of distinguished features forming a feature vector. Each feature is associated to a class by employing training data and a classifier is said to be "trained". Consequently, this trained classifier is then fed with the test data to classify each class accordingly. Normally, solution offered by LB method is optimal in a sense that it reduces the probability of misclassification. However, FB method is more famous in practical scenarios mainly due to its robustness to mismatch caused by channel such as frequency offset or timing error [1].

FB methods can be broken down into two main sub systems: Feature extraction sub system and Classifier sub system. This is shown in Fig. 1.

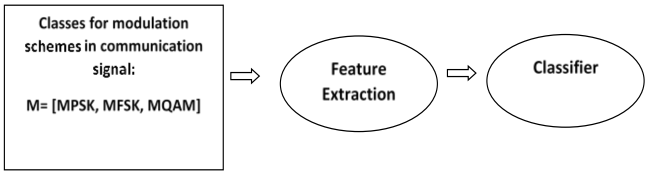

Fig. 1. Feature based classification subsystems.

In the first subsystem, features are first extracted from the received signal. Some common features that has been used in the previous work are: instantaneous frequency [6], higher order cumulants [7]-[14] and time-frequency domain analysis [8]. Second subsystem comprises of a designed classifier for the input training and test signals. These include Support vector machines (SVM), multiclass one vs all (OVA) SVM, K-nearest neighbor (KNN) classifier, Nayes Bayes based classifier, tree based classifiers and regression based classifier like AdaBoost, Bag etc.

One of the most important parts of the FB based classification is selection of a feature vector from within the generated features. Feature selection can be further broken into two main components: filter based and wrapper based. Filter based algorithms are generally used as a preprocessing step before applying the wrapper based filter selection. Some common filter based approaches are *p*-value based t-test statistics, misclassification based filtering. On the other hand, Sequential feature selection, Principle component analysis based dimension reduction or Singular value decomposition based feature selection are some of the wrapper based methods.

In this paper, the aim has been to design a feature selection algorithm for 3 groups of digitally modulated signals mainly PSK, FSK and ASK. The features in Ambiguity function (*AF*) domain of the received signal have been used [15], [16]. Different modulation formats have distinct features in *AF* domain [9]. The signal received has been firstly converted to *AF* domain. Selection of features have been done in two steps: 1) using t-test statistics for preprocessing 2) SVD based selection, with dominant Eigen vectors, to form final feature vector f_d. Following this, feature vector f_d is fed to second subsystem comprising of classifier. Performance of two classifiers has been investigated mainly: multiclass one vs all (OVA) Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifier.

The paper is organized as follows. System model is defined in Section II. In Section III, methods for feature generation and selection are discussed respectively. Section IV gives the details about designing of the two classifiers.

Manuscript received March 20, 2017; revised May 22, 2017.

Afan Ali and Fan Yangyu are with School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China (e-mail: afanali85@yahoo.com, fan_yangyu@nwpu.edu.cn).

DOI: 10.7763/IJCTE.2017.V9.1174

Section V discusses simulation results.

II. System Model

The received base band signal is given by

$$r(t) = s(t) + n(t) \tag{1}$$

where $n(t)$ is Additive White Gaussian Noise (AWGN) and s(t) is dependent on modulation type.

The expressions used for the received sequence is given by

$$s_{PSK} = A\mathrm{Re}[\sum_{K} C_K e^{j2\pi f_c t} g(t - kT_s)]$$
$$C_k = e^{j\frac{2\pi i}{M}} \tag{2}$$

$$s_{FSK} = A\mathrm{Re}[\sum_{K} e^{j2\pi(f_c + \Delta f_k)t} g(t - kT_s)] \tag{3}$$

$$s_{QAM} = A\mathrm{Re}[\sum_{K} C_K e^{j2\pi f_c t} g(t - kT_s)]$$
$$C_k = a_k + jb_k, a_k, b_k = 2i - M - 1 \tag{4}$$

where s_{PSK} =PSK modulated received sequence, s_{FSK} =FSK modulated received sequence, s_{QAM} =QAM modulated received sequence, i=0,1,2,……,M-1, A is the power of the received signal, C_K map the transmitted symbols, Ts is the symbol period, f_c is the carrier frequency, M is the modulation level and $g(t)$ is the finite energy signal with a T_s duration.

III. Features Generation and Selection

Original data is a digitally modulated signal categorized in one of the following M classes:

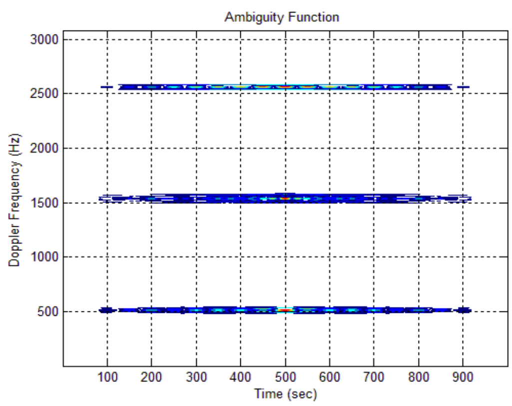

Fig. 2. Ambiguity Function (AF) of three groups.

M= [MPSK, MFSK, MQAM]. The received sequence for each of the 3 classes has been formed into v_1, v_2 and v_3 matrices. For each received signal matrix, a 2-D ambiguity function $AF_v^{M \times N}$ (M and N are frequency and time sizes respectively) has been computed as shown in Fig. 2. Rows of this ambiguity function matrix are the observations and columns represents features. Some features may be redundant and therefore, further processed to compute to get the optimum feature vector, v_D.

A. T-test Statistics

Firstly, the computed $AF_v^{M \times N}$ for each class has been combined together by connecting the columns and a high dimensional vector, v_L formed. T-test has been performed on v_L to compute the p-values. This is shown in figure 3.

Fig. 3 shows the empirical cumulative distribution function of p-values in order to give a general idea about how well-separated the 3 groups are by each other. There are about 35% of features having p-value equal to zero and almost 50% of features with p-values smaller than 0.05, meaning there are around 50% of the features that have strong discrimination power. These features can be then sorted according to their p-values (or the absolute values of the t-statistic) and selected from the sorted list. This vector is denoted by v_P.

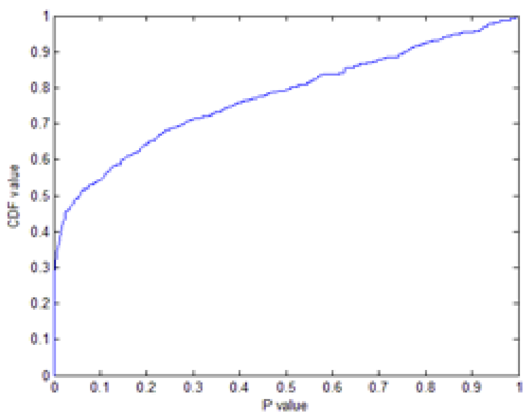

Fig. 3. T-test statistics.

B. SVD Based Algorithm

Now, Singular Value Decomposition (SVD) based algorithm has been applied with vector v_P as input.

Table I shows the feature selection methodology adopted.

TABLE I: SVD Based Feature Selection Algorithm

i. Preprocessed feature vector v_p is input.
ii. Then, compute the centralized data.
iii. Compute SVD to get the principle components.
iv. Get the dimensions having most of the variation. Therefore, only the dominant eigenvectors are selected, e.g. representing the 95% of the data.
v. Then the leverage scores are computed using Eigen values of the principal components. That is, we take the norm of the Eigen vector's coefficients.
vi. Then, the leverage scores are sorted in descending order.
vii. Check 20-50 indices of the vectors with the largest leverage scores.
viii. Number of features with best correct classification is chosen as the final feature vector v_F.

This procedure is the same as the column select problem discussed in [10]-[17]. Firstly, the leverage scores for each dimension has been computed and then features ranging from 20 to 50 have been selected from them and tested on classifier to see which one gives the best correct classification. They are selected based on largest scores as weights.

IV. Multiclass SVM and KNN Classifier

SVM classifier is inherently two-class classifier. It has been modified for multiclass classification in the project. One vs all (OVA) technique has been used. Gaussian radial basis function (rbf) kernel with scaling factor, $\sigma = 1$ has been used. Kernel function basically maps the training data into kernel space. Least square ('LS') method has been used to find the separating hyperplanes.

KNN classifier is both simple and fast. It is designed using 4 nearest neighbor (k=4). Distance metric used in the designing is Euclidean.

V. Simulation and Results

Table II shows the simulation parameters set for the modulated generated data signals.

TABLE II: Simulation Parameter Setting

Parameter/Signals	PSK	FSK	QAM
Sampling frequency f_s (kHz)	10	10	10
Symbol Rate (per sec)	100	100	100
No. of symbols	20	20	20
Modulation Order	2 ,4, 8,16	2 ,4, 8,16	2 ,4, 8,16

In the first phase of the experiment, rigorous cross validation test is applied to the feature selection algorithm and classifier using training data only. Batches ranging from 10%-90% of the training data are used respectively to train the classifier and the rest is held back to test the trained classifier. Elements in feature vector are varied to get optimum feature vector. There are two main scenarios in this experiment mainly:

Scenario 1: 100 features from t-test statistics and 20, 30, 40 features from SVD based algorithm shown in figure 4.

Scenario 2: 500 features from t-test statistics and 20, 30, 40 features from SVD based algorithm shown in Fig. 5.

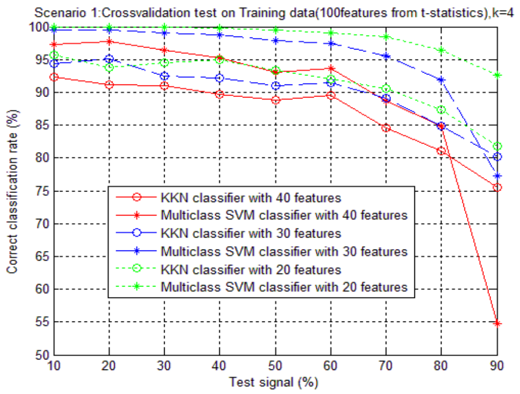

Fig. 4. Cross validation test on training data, Scenario 1.

As can be seen from Fig. 4, multiclass SVM classifier with 20 best selected features outperforms other combinations in the validation test. It achieves a 100 % correct classification rate when almost 50% training sequence is used to train the classifier.

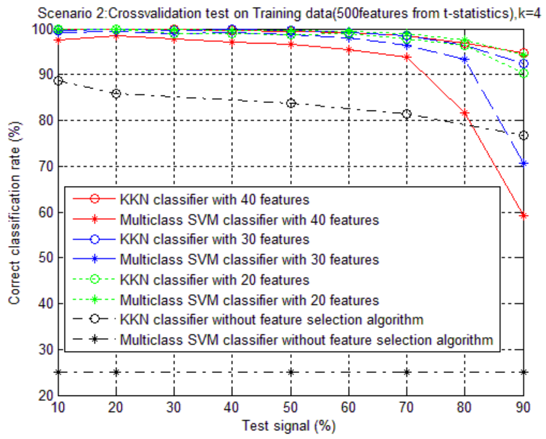

Fig. 5. Cross validation test on training data, Scenario 2.

From Fig. 5, with 50% features from t-statistics test, it is evident that KNN classifier performs slightly better than multiclass SVM esp. with 30 and 40 features selected from SVD based algorithm. However, best performance is still achieved by multiclass SVD classifier with 20 features selected from SVD algorithm. Result for correct classification with no feature selection algorithm applied to the training data is also depicted in this graph. It can be seen that without feature selection algorithm, both the classifiers under performed.

In the second phase, test data is classified using training samples to train the two classifiers. Figure 6 shows this result. It is evident from the graph that two thick lines representing multiclass SVM with 20 features and KNN classifier with 40 features shows the best result. This was expected after the cross validation result in the previous experiment. However, results also show that multiclass SVM shows better results than KNN classifier even when data contains 90% of test samples.

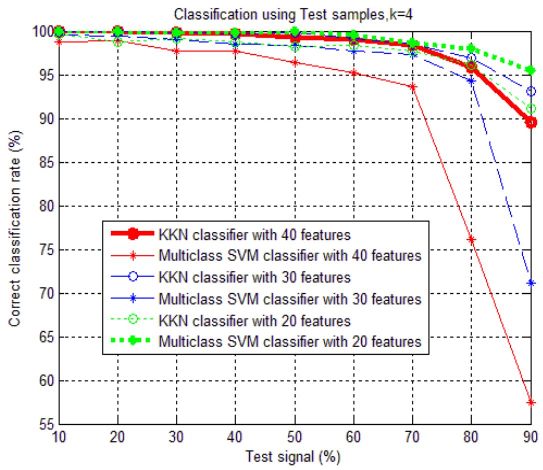

Fig. 6. Correction classification with test samples.

VI. Conclusion

This paper investigates the automatic modulation classification using pattern recognition approach. Features are generated using ambiguity function domain of the modulated signal. This is because different modulated signals have distinct characteristics in ambiguity function domain. Non-redundant, efficient and robust features are selected from amongst the generated features using filter and wrapper based methods. Preprocessing of features has been done via filter method known as t-statistics test. These preprocessed features are then input to a wrapper based method which selects the most dominant features using Singular Value Decomposition (SVD). Classification is then done using two classifiers mainly KNN classifier and multiclass One vs All (OVA) Support vector machine (SVM). Different ratios of test and training signals with different number of selected features have been used in simulation to investigate the two classifiers performance. Desired correct classification rate is achieved using both the classifiers with 20, 30 and 40 features in the feature vector. However, multiclass SVM outperforms KNN classifier in most cases.

This methodology can work in large number of modulation groups under low Signal-to Noise (SNR) ratio as noise is centralized in the ambiguity function. Future work can consider the performance of the system in low SNRs. A lot of new features for modulated signal like high order cumulants, invariant moments and cyclostationary features can also be combined with the ambiguity function matrix to make the designed system more robust.

REFERENCES

[1] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "A survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Comm.*, vol. 1, pp. 137-156, 2007.

[2] Z. C. Zhu and A. K. Nandi, *Automatic Modulation Classification: Principles, Algorithms and Applications*, Wiley, 2014, ch. 1, pp. 25-34.

[3] V. Iglesias and J. Grajal, "Automatic modulation classifier for military applications," in *Proc. Signal Processing Conference, 19th European*, 2011, pp. 1814-1818.

[4] J. L. Xu, W. Su, and Z. Mengchu, "Likelihood-ratio approaches to automatic modulation classification," *IEEE Trans. Syst., Man, and Cybernetics. Part C: Appl. and Reviews,* vol. 41, pp. 455-469, 2011.

[5] O. A. Dobre and F. Hameed, "Likelihood-based algorithms for linear digital modulation classification in fading channels," in *Proc. IEEE CCECE,* 2006.

[6] P. Panagiotou, A. Anastasoupoulos, and A. Polydoros, "Likelihood ratio tests for modulation classification," in *Proc. IEEE MILCOM,* 2000, pp. 670- 674.

[7] H. Sarieddeen, M. M. Mansour, M. A. Louay Jalloul, and A. Chehab, "Likelihood-based modulation classification for MU-MIMO systems," in *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP),* Dec. 2015, pp. 14-16.

[8] A. Hazza, M. Shoaib, S. A. Alshebeili, and A. Fahad, "An overview of feature-based methods for digital modulation classification," in *Proc. 1st International conference on Communications, Signal Processing and their Applications (ICCSPA),* 2013.

[9] H. Deng, M. Doroslovacki, H. Mustafa, J. Xu, and S. Koo, "Instantaneous feature based algorithm for HF digital modulation classification," in *Proc. CISS Conf., 2003, John Hopkins Universit*y.

[10] Y. Yang and S. S. Soliman, "Statistical moments based classifier for MPSK signals," in *Proc. GLOBECOM,* 1991, pp. 72-76.

[11] E. E. Azzouz and A. K. Nandi, "Automatic modulation recognition of communication signals". *Kluwer Academic,* 1996.

[12] K. C. Ho, W. Prokopiw, and Y. T. Chan, "Modulation identification by the wavelet transform," in *Proc. IEEE MILCOM,* 1995, pp. 886-890.

[13] M. L. D. Wong and A. K. Nandi, "Efficacies of selected blind modulation type detection methods for adaptive OFDM systems," *EURASIP JASP*, 2009.

[14] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun*., vol. 48, pp. 416-429, 2000.

[15] Z. M. Xu, G. A. Bi, and X. M. Li, "On application of time-frequency analysis to communication signal detection and estimation," in *Proc. 9th IEEE International Conference on Control and Automation (ICCA)*, 2011, pp.1044-1048.

[16] H. J. Zhang, G. A Bi, S. G. Razul, and C. M. S. See, "Supervised modulation classification based on ambiguity function image and invariant moments," in *Proc. 9th IEEE Conference on Industrial Electronics and Applications,* 2014, pp. 1461-1465.

[17] M. W. Mahoney and P. Drineas, "CUR matrix decompositions for improved data analysis," in *Proc. the National Academy of Sciences,* February 2009.

Afan Ali is a Ph.D. student in the School of Electronics and Information, Northwestern Polytechnical University. He received his master's degree in electronics and communication engineering from Australian National University, Australia in 2011. He completed his B.S. degree in telecommunication engineering from National University of Computer and Emerging Sciences, Pakistan in 2008. His research interests include pattern recognition, digital communications and signal processing.

Fan Yangyu is a professor in Communication and Information Engineering Department at Northwestern Polytechnical University. He is also a director of the Institute of Signal Processing and Wireless Optical Communications. He completed his Ph.D degree from North Western Polytechnical University in 1999. He completed his MSc. and BSc. from Shaanxi University of Science & Technology in 1992 and 1982 respectively. He is also administrative director of Shaanxi Society of Signal Processing. His research interests include signal processing, wireless optical communications, image processing and virtual reality.