Classification of Underwater Acoustic Signals Using Multi-Classifiers

Khizer Feroze[1], Sidra Sultan[1], Salman Shahid[2], Faran Mahmood[1]
[1]Department of Electrical Engineering, Institute of Space Technology, Islamabad, Pakistan
[2]Teradata, Pakistan

***Abstract*— Recent advancements in maritime technologies, pure oceanic research and military techniques related to underwater acoustics have introduced us to a new class of problems in digital signal processing. Apart from acoustic measurements, underwater classification to identify signatures poses a challenging task at hand due to typical profile of anthropogenic noise which is very complex. Moreover, unavailability of large relevant datasets makes it hard to achieve high accuracy in classification. Underwater acoustic channel is a dynamic one that not only includes noise due to fish and marine life but also that from commercial human activity and military operations. This work shows that a single classification approach is not enough to classify signals with very high accuracy. Several techniques of underwater acoustic classification using objects' unique acoustic signature have been investigated. This classification process using acoustic signatures is divided into two stages. In the first stage, preprocessing and feature extraction needs to be performed on the given dataset - which typically includes a variety of marine life and anthropogenic signals. In the classification stage, multiple techniques including Naive Bayesian, K nearest neighbors, Artificial neural networks, Support vector machines and Hidden Markov Models have been performed to evaluate classification results. The shortcomings of each classifier are discussed thoroughly in this paper. The main challenges for the robust and fast classification of underwater signals are also described to mitigate them in future. Moreover, open research issues are also discussed, and possible solution approaches are outlined.**

Keywords— Acoustic signature, robust classifier, underwater signals, feature extraction, K nearest neighbor, Naïve bayesian

I. INTRODUCTION

Whenever a song is playing on radio, people can easily guess about the singer through simple recognition. The human recognition model simply includes ears, cochlea and a central nervous system including brain and nerves. Every person can be identified through its unique voice, thus making it a significant characteristic feature for identification and recognition purpose. Underwater medium consists of a variable number of signals emitted from fishes and submarines which includes machines to perform monitoring tasks over a given area for the core security purposes as per military requirements. Every fish has its voice, likewise, every ship produces a unique combination of its machine and propeller sounds, thus making a unique "Acoustic Signature" for identification [1-6]. In the recent years, due to advancements in oceanic research and technology, the classification of underwater acoustic signals has gained much attention. Increasing levels of certain activities of humans at the sea have caused much growth in noise level. Those activities include marine and transport activities etc. Unfortunately, the propagation and anthropogenic noise has heavily distorted "acoustic signature" [7]. Due to these increased anthropogenic noise, the classification problem of underwater targets has posed a great problem in this area for researchers.

II. UNDERWATER ACOUSTIC SIGNAL CLASSIFICATION AND RECOGNITION (UASR)

A. Literature Survey

Although similar work exists in the prior art but our work on the underwater signals can be compared to some other models developed in the prior art. In most of the cases, the audio signals (whose medium is air) are processed. Work on the underwater signals covers mostly the sound of mammals and do not incorporate anthropogenic sounds.

The work of {Martí Farriol, 2012} classification is done mostly on dolphin whistles and artificial sonar signals. All the underwater acoustic signals are not being covered, instead work is done only on dolphin whistles and self-generated sonar signals. So, the main difference is in data type as compared to our work where we have included all underwater data types (Ships, submarines, vessels etc.).

Moreover, the data used in their model is taken from one environment only. The platform that they have used in ANTARES, which is an underwater platform in the Mediterranean Sea. The conditions stay same for all the underwater signals ever recorded. In our work we have developed a generic model where the sound can be classified regardless of the environment from where it is coming. We have taken the underwater channel and incorporated all the effects that can take place. For testing of this generic model, we have developed a noise generator that distorts the signal exactly according to underwater environments.

In their work {Pohjalainen, 2007} writer has not specified through tests that how much noise immunity can be tackled by the classifier. But in our work, we have specified the noise immunity level of each feature, and classifier.

The algorithms that used in their model are entirely different from ours. However, this work served as a good base for getting started with UWA classification.

Moreover, Research work can be divided into two types: Firstly, Feature extraction techniques and secondly their identification mechanisms. Although, many techniques have been adopted by researchers for feature extraction purpose, however it is observed that many features which belong to multiple classes are overlapped mainly because of the vessel

978-1-5386-3564-3/18/$31.00 ©2018 IEEE

radiated noise which makes it complicated [4]. To mitigate vessel radiated noise, several classification approaches are proposed including back propagation neural network [2, 3] and probabilistic neural networks [4]. Kohonen Neural Networks is also used as a classifier for vessel radiated noise [6]. Further in the research work, it is observed that same methods can be deployed in many other applications for civilian as well as military purposes [4]. However, in every application, the underwater acoustic signal should be preprocessed after recording and trained by any classifier to classify the signals into well-defined classes.

Many techniques for signal processing have been proposed for the extraction of targets submerged in sea/ocean, for their identification purpose. One of which is Fourier transform. However, this transform poses difficulties for analyzing short-time transient sounds. For compensating the problem, assorted short time Fourier transforms (STFT) are being developed by utilizing windows [6]. Several alternatives to assorted STFT are also being proposed. One of which is proposed is Cohen, who developed a time frequency distribution through producing the spectrums at alternate time intervals [2].

B. Problem Statement

Due to increased anthropogenic noise, the classification of Underwater acoustic signals cannot produce efficient results by deploying denoising mechanisms mainly because of the pink noise, additive white Gaussian noise (AWGN) and the impulsive noise which remains present [8, 9]. It is difficult to accurately model each and every source of noise. The most effective method for underwater acoustic signal classification is through feature extraction technique. Those features are unique depending on the unique acoustic signature of a particular ship or fish. Feature extraction method performs feature vectorization by converting the discrete signals into a sequence of feature vectors [10]. These feature vectors are then trained through a robust classifier [11]. It is observed that a single feature cannot train the classifier efficiently. Hence multiple features including time-domain and frequency domain signals are extracted.

The primary goal behind this research is that a single classification approach is not enough to classify signals with much accuracy. Several classification techniques of underwater acoustic signals through its unique acoustic signature should be investigated. Each type of signature has distinct characteristics and should therefore be classified through a robust classifier.

III. PROPOSED WORK

The proposed scheme is to deploy multiple classifiers. Following steps explain the proposed model.

1) Identification of Classification & Accumulation of Data repository

For the experiment, these includes the mammal noise (whale, dolphin etc.) and shipping noise (commercial, fishing, submarines etc.) Sound files falling in the identified categories are accumulated from online repositories and existing resources.

2) Vectorization of Categorized sounds

The features identified as part of previous work (Automatic Classification of Underwater Acoustic Signals) are further optimized to be discriminative as possible between considered classes and extracted from the categorized files using relevant signal processing algorithms. The features identified are further tailored for automated discrimination/isolation of unique sound sources within accumulated data recordings.

3) Development and Testing of Classification Algorithms

The classifying algorithm chosen is dependent upon the level of discrimination achieved in the feature identification and extraction phase. The classifiers chosen are K nearest Neighbors [13][14], Support Vector Machines [15][16], Artificial Neural Networks [3][18][19], Naïve Bayesian [17][12] and Hidden Markov Models [20][21] .

IV. IMPLEMENTATION

A. Preprocessing of signals

The sounds which are in very small amount and recorded with single hydrophone at same depth, those types of sounds are divided into two equal parts. We have taken 50% as training and other 50% is fed to noise generator that immerges noise into it and makes its correlation function between 65 to 75%.

B. Feature Extraction

The classification model starts with the feature extraction part where signals are manipulated, and useful trends are extracted from them. Features comprise of three categories: time domain, frequency domain and perceptual features.

1) Time domain features

Following time domain features are being used in our model.

a) Short Time Energy

STE is used for region extraction purpose, in which regions containing audio content and the noise can be identified. The noise is further rejected.

b) Root mean square

The same region extraction algorithm is developed through root mean square with a few refinements.

c) Zero Crossing

Zero crossing rate (ZCR) is used to differentiate between different speakers and between different instruments. It determines the number of times a sound passes through a zero level.

d) Gradient Index

Gradient index serves in separation of noisy and noise free parts. Gradient index (GIA) helps in developing a special trend setting in the frames containing actual content.

e) Unit Lag

Unit-lag normalized autocorrelation is computed using the autocorrelation function. Autocorrelation is the correlation of a signal, matched with its own delayed version.

2) Frequency domain features

Frequency domain features are used to manipulate a sound by taking Fourier transform.

a) Spectral Flatness

Rapid fluctuations within a signal is computed by spectral flatness. The first use of this feature is the separations of noise from noise free regions. The regions containing noise components only will show less variations (will be flat) in the spectral flatness trend, while the acoustic regions will have more variations and less flatness.

This feature is incorporated in our work to check the noise immunity as most of the signals are noisy.

b) Spectral Roll off

Frequencies with in a frame, below which 85% of the magnitude distortion are depicted through this feature. Though it is a simple feature to analyze, yet it has got some good noise immunity. Spectral roll off is mostly not used in audio signals classification (where the medium is air), however employing this feature to underwater environment can suffice the noise immunity requirements to certain extent.

3) Perceptual features

a) Mel Frequency Cepstral coefficient

MFCC is a cepstral feature that works on human ear perception. It makes use of cepstral features.

In our work, we were provided with all the three categories of those features, so they were being utilized in classifiers and their effectiveness and noise immunity from research point of view are studied. The main working of our system is shown through block diagram in fig 2.Multiple Classification Tecniques

The work of classifier is to create classes and assign the corresponding class to the sound signal depending upon its feature vectors [12]. This correspondence is done through mapping and training the dataset. Here, in this research paper we have focused on multiple classification algorithms. The best classier that works well in underwater noisy environment is selected.

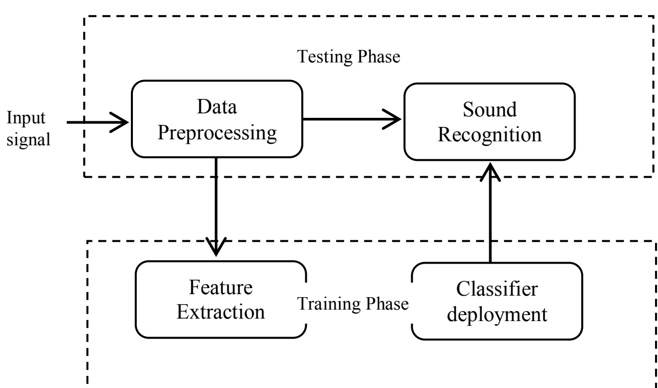

Fig. 2. Block diagram of Under-water acoustic recognition system

V. Experiments and Results

A. Dataset

Data contains 29 machines and 24 fishes' sounds (see table 1). Each having different lengths and sounds recorded in different oceans. The same sounds which are taken at different oceans and on different depth are directly taken 50% as training and 50% as testing.

The sounds which are in very small amount and recorded with single hydrophone at same depth. We split that type of sounds into two equal parts and took 50% as training and other 50% were given to noise generator that immerged noise into it and made its correlation function between 65 to 75%.

B. Noise Generator

The dataset for our work is being taken in ideal conditions where hydrophones record the sounds. The noise generator incorporated the underwater parameters like speed of air, depth etc. to see their effects on the signals.

While a signal travels through the underwater environment, it gets itself immersed with certain type of distortions.

If we are using a lower central frequency, and the attenuation occurs due to the distance, the signal become more prone to the noise. Signal to noise ratio decreases and classification become more challenging. Moreover, wind speed also effects the signals. These distortions are being modeled by developing a noise generator as shown in fig 4. Since, for this work, we did not need to develop a noise generator as it is already being available for our work.

The work of this noise generator is to induce controlled noise in the signals in order to test the condition under which the classifier works best for the noisy signals.

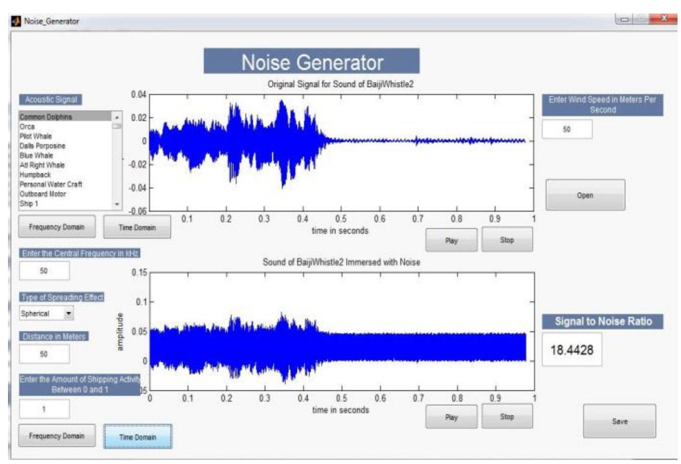

Fig. 4. Working of a Noise Generator

C. Experiments

To classify that the test sound is either from fish or machine dataset, we extracted features of the test sounds and fed them to each classifier. After getting the best classifier that separates fish sound or machine sound, we further classified them based on the feature set to determine the type of machine or fish.

There are total 2^8=256 possible combinations of features. For example, if we take root mean square and short time energy then it is one possible combination and if we take root mean square and zero crossing rate then this will be another possible combination. So, we firstly found each possible combination of feature and gave them separately to five different classifiers. Thus, our working is based on five experiments where in each experiment, the parameters of noise remained the same as follows.

Wind speed = 20 m/s
Central Frequency = 100 kHz
Distance = 1000 m

In the first set of experiments, we developed the K nearest neighbor classifier by training the signals obtained from dataset. Different values of K are taken for different combinations of feature set. Then the system is tested against the classifier for test signals in which noise has been incorporated in a controlled environment.

In the second set of experiments, we developed support vector machines (SVM) model. SVM works for 2 classes only, but we have developed an algorithm that works for multiple classes.

Kernels are useful for optimization when there are lot of features used. Here, we have used kernel method to get the most efficient result. We have taken following SVM kernels to classify data.

- Linear
- Quadratic
- Polynomial

In the third set of experiments, naïve Bayesian model is used with the same noise parameters being incorporated.

In the fourth and fifth set, Artificial Neural Networks and Hidden Markov models are being used to develop a robust classifier respectively.

Table 1: Fish and machine dataset

Fishes	Machines
Blue_whale	Airgun
Bowhead whale	ATOC transmission
Brydes whale	IceBreaker
Fin whale	Bubble Curtain
Gray whale	Dredging
Humpback whale	Ocean Acoustic Tomography T
Minke whale	Outborad Motor
Right whale	Pile Diving
Southern whale	Personal Watercraft
Crabeater seal	Cruise Ship
Leopard seal	Survelliance towards ASSL
Sperm whale	Surtass LFA sonar
Bottlenose dolphin	Torpedo
Spinner	Underwater Breathing Apparatus
Mealon headed whale	Wind Turbines
Pilot	Fighter Nearing Dock
Gray	Gotlands Turbines
Rissos	High Speed Propeller
SeaRobbin	Hydro Thermal Vent
Trumpet	Large Commerical Ship
Weak Fish	Low Frequency Active Sonar
White sided dolphins	Merchant Vessel
Harp Seals	Mid Sized Vessel
Winged Seals	Outboard 60Hp
	Personal Water craft
	Samsoe Turbines
	Small Diesel
	Vindeby Turbines
	Whinning Propeller

D. *Results*

Experimental results are described below, by using each possible combination of features.

1) K nearest neighbour

It is showed from the output results that the truly classified rate is 63.7% for machine dataset (see table 2). For the sake of simplicity, only a few results are shown. The optimum features used are root mean square and zero crossing rate. Root mean square is low level feature due to which it works better with the simple classifier like KNN.

The results showed that by using time domain features, we get the best results instead of using frequency domain features for machine dataset.

Moreover, the identification rate is 78.2% for fish dataset by using the MFCC and unit lag as features. The significance of MFCC lies in Mel-frequency filters that are non-linear triangular filters.

Hence, frequency domain features are efficient for classification of fishes.

Table 2: Classification rate for KNN

KNN (value of k)	Features	Accuracy
9	All	23.539%
15	All	30.88%
3	RMS, ZeroCrossing	63.67%
3	MFCC, UnitLag	78.2%

2) Support Vector Machines

In the second set of experiments, it is noted from the results that SVM works efficiently with polynomial kernel for both machines and fishes (see table 3). It gives 82.5% accuracy for fish dataset and 68.32% for machine dataset.

Table 3: Classification rate for SVM

Kernel	Features	Accuracy
Linear	SpectralRolloff, MFCC	55.31%
Linear	STE, SpectralFlatness	78.73%
Quadratic	RMS, STE, ZeroCrossing, GradientIndex, Spectral Flatness	67.8%
Quadratic	SpectralRolloff	80.14%
Polynomial	ZeroCrossing	82.5%
Polynomial	SpectralFlatness, SpectralRolloff	68.32%

3) Naïve Bayesian approach

In the third set of experiments, analysis from the results showed that Naïve Bayesian gives an average result of 55.31% for machine dataset with unit lag and spectral flatness features while for fish dataset it performs better than the other classifiers with accuracy of 81.67% (see table 4).

The spectral flatness feature results in more randomness in case of white noise and less randomness in case of pink or impulse noise. Therefore, the type of noise changes the accuracy rate.

Thus, it is concluded that naïve Bayesian does not work well for machines data. Moreover, none of the time domain or frequency domain features work best for the data set. Thus, naïve Bayesian cannot be used alone for classification.

Table 4: Classification rate for Naïve bayesian

Features	Accuracy
SpectralRolloff, MFCC	55.31%
RMS, STE, ZeroCrossing, Spectral Flatness	81.67%

4) Artificial neural networks

This approach is giving the accuracy rate of 67.84% for machine dataset and for fish dataset, it is 79.9 %. Only the best results are being shown. (see table 5).

Table 5: Classification rate for ANN

Features	Accuracy
STE, Mel Frequency cepstral	79.9%
STE, Spectral Roll off	67.84%

5) Hidden Markov models

Hidden Markov's model which is broadly using in speech recognition is giving equal results for both classes of data that is 78 % (see table 6).

Table 6: Classification rate for HMM

Features	Accuracy
Unit Lag SpectralRolloff, MFCC	78.25%
SpectralRolloff, Spectral Flatness, Unit Lag, MFCC	78.72%

The results conclude that Hidden Markov models works best for machines if we take time domain features. Although time domain features do not really work with noisy signals but with HMM classifier, the noisy part is mitigated efficiently.

Moreover, HMM does not work well with sounds of less duration. Hence HMM needs a huge database for training.

VI. Conclusion

In this paper, we implemented multiple classification based identification techniques to classify underwater acoustic signals under controlled noise environment. The noise parameters for each type of classifiers were kept same. The resultant model gave promising results in terms of accuracy.

We used dataset of 29 machines and 24 fishes. We found by several experiments that for correct classification, same feature set cannot be used. Zero crossing is a dominant feature for machines while MFCC should be considered for fishes. Hence, Hidden Markov's Model and support vector machines with polynomial as a kernel would be a very good choice for classification of Machine data and Fish data respectively.

However, SVM is somehow slower than other classifiers so Naïve Bayesian can be used for speedy output.

VII. FUTURE WORK

In the future, some hybrid approaches can be deployed to get more efficient results. Combination of Classifiers instead of single one along with the denoising approach in preprocessing phase can provide much better results. Moreover, feature weighing algorithms like Ada Boost can also help in achieving promising results. In addition, more features can be extracted to study their effectiveness on the dataset.

ACKNOWLEDGMENT

We express our profound gratitude to Dr. Salman Shahid, Mr. Faran Mehmood and Mr. Abdullah Iqbal for their guidance and encouragement throughout this research paper.

REFERENCES

[1] P. Chestnut, H. Landsman, and R. W. Floyd, "A sonar target recognition experiment," *The Journal of the Acoustical Society of America,* vol. 66, no. 1, pp. 140-147, 1979.

[2] M. R. Azimi-Sadjadi, D. Yao, Q. Huang, and G. J. Dobeck, "Underwater target classification using wavelet packets and neural networks," *IEEE Transactions on Neural Networks,* vol. 11, no. 3, pp. 784-794, 2000.

[3] C.-H. Chen, J.-D. Lee, and M.-C. Lin, "Classification of underwater signals using neural networks," *淡江理工學刊,* vol. 3, no. 1, pp. 31-48, 2000.

[4] R. Chen, H. Li, and S. Tang, "Association rules enhanced classification of underwater acoustic signal," in *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, 2001, pp. 582-583: IEEE.

[5] H. F. Hashem, "Automatic classification of underwater sonar signals," in *Neural Network Applications in Electrical Engineering, 2004. NEUREL 2004. 2004 7th Seminar on*, 2004, pp. 121-125: IEEE.

[6] V. Lobo, "Ship noise classification: a contribution to prototype based classifier design," *Departamento de Informatica. Lisbon, Universidade Nova de Lisboa,* 2002.

[7] V. Lobo and F. M. Pires, "Ship noise classification using {K} ohonen networks," 1995.

[8] J. Martí Farriol, "Underwater signals classification and clustering," Universitat Politècnica de Catalunya, 2012.

[9] M. Yuanliang, "Challenges and potential of development for current signal processing on underwater acoustic," in *the National Acoustics Conference*, 2002, pp. 5-8.

[10] J. Pohjalainen, "Methods of automatic audio content classification," 2007.

[11] S.-c. Li, D.-s. Yang, and L.-p. Jin, "Classifying ships by their acoustic signals with a cross-bispectrum algorithm and a radial basis function neural network," *Journal of Marine Science and Application,* vol. 8, no. 1, pp. 53-57, 2009.

[12] M. Kantardzic, *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons, 2011.

[13] P. Viswanath and T. H. Sarma, "An improvement to k-nearest neighbor classifier," in *Recent Advances in Intelligent Computational Systems (RAICS), 2011 IEEE*, 2011, pp. 227-231: IEEE.

[14] S. Sun and R. Huang, "An adaptive k-nearest neighbor algorithm," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference on*, 2010, vol. 1, pp. 91-94: IEEE.

[15] Z. Xuegong, "Introduction to statistical learning theory and support vector machines," *Acta Automatica Sinica,* vol. 26, no. 1, pp. 32-42, 2000.

[16] J. Hu, D. Li, Q. Duan, Y. Han, G. Chen, and X. Si, "Fish species classification by color, texture and multi-class support vector machine using computer vision," *Computers and electronics in agriculture,* vol. 88, pp. 133-140, 2012.

[17] K. M. Leung, "Naive bayesian classifier," *Polytechnic University Department of Computer Science/Finance and Risk Engineering,* 2007.

[18] T. L. Hemminger and Y.-H. Pao, "Detection and classification of underwater acoustic transients using neural networks," *IEEE Transactions on neural networks,* vol. 5, no. 5, pp. 712-718, 1994.

[19] J. Ghosh, L. Deuser, and S. D. Beck, "A neural network based hybrid system for detection, characterization, and classification of short-duration oceanic signals," *IEEE Journal of Oceanic Engineering,* vol. 17, no. 4, pp. 351-363, 1992.

[20] L. Rabiner and B. Juang, "An introduction to hidden Markov models," *ieee assp magazine,* vol. 3, no. 1, pp. 4-16, 1986.

[21] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE,* vol. 77, no. 2, pp. 257-286, 1989.