Classification of Underwater Acoustic Signals Using Various Extraction methods

N. Korany[1], M. Abd Elzaher und H. Khater

E. E. Dept., Faculty of engineering, Alexandria University, Egypt. E-Mail: nokorany@hotmail.com[1]

Abstract

Automatic classification of underwater acoustic signals (ACUS) is used to enable a navy to identify the ships by recognizing the underwater sound that they produce. In this paper, three types of features, Mel Frequency Cepstral Coefficients (MFCC); Perceptual Linear Predictive Cepstral Coefficients (PLPCC) and Relative Spectral Perceptual Linear Predictive Cepstral Coefficients (RASTA-PLPCC) are extracted for the classification problem. The classifier identification rate is calculated for each type of extracted feature using Gaussian mixture model (GMM). The calculation is repeated while varying the number of coefficients for each type, and the performance of the recognition model is investigated.

Introduction

ACUS is considered as a complex problem because of the large variability in both temporal and spectral characteristics in signals even obtained from the same source. When designing a classifier, usually it is not known priori which features are the best for the classification problem. The best solution is to generate more than one feature and then choose the best.

In this paper statistical identification technique using GMM was implemented[1]. MFCC, PLPCC and RASTA-PLPCC are used for the classification problem. The performance of the recognition process is evaluated while varying the number of coefficients for each feature type. In all, sound must be recorded, pre-processed, the features are extracted and classified into one of series of possible classes.

Feature Extraction

In this paper, *MFCC*, *PLPCC and RASTA-PLPCC* are extracted and are used to describe the sonar signal.

MFCC

For each frame of the signal, MFCC are computed as follows[2].

a) Obtain the amplitude spectrum
b) Warp to Mel (a perceptually-based) scale.
c) Sum the contents of each band.
d) Compute the logarithm of each sum.
e) Compute the DCT of the bands.
f) Compute a number of coefficients of a certain order using the cosine transform.

PLPCC

The PLP model simulates the properties of hearing by estimating the critical-band spectral resolution, the equal-loudness curve and the intensity-loudness power law, and then the auditory spectrum is obtained. Finally, the autoregressive coefficients are solved. The PLPCC are computed as follows[3].

a) Critical-Band analysis for the sound signal.
b) Equal loudness pre-emphasis.
c) Intensity-Loudness conversion.
d) Inverse Discrete Fourier Transform.
e) Solution for the autoregressive coefficients of the all-pole model.
f) Extract the cepstral coefficients from autoregressive coefficients.

RASTA-PLPCC

The underlying principle of RASTA processing is to suppress any constant component in each frequency channel of the short term auditory spectrum prior to the all-pole model. This is done by filtering the time trajectories of each component of the spectral representation of the signal [4]. The transfer function of the filter is given in equation (1).

$$H(z) = 0.1z^4 \frac{2 + z^{-1} - z^{-3} - 2z^{-4}}{1 - 0.98z^{-1}} \qquad (1)$$

Database

Ships sounds with predefined characteristics like number of shafts , blades per shaft, speed, direction and distance are simulated. The database consists of sound signals that are produced by three-surface and three-subsurface ships with different numbers of blades and shafts. Every ship's sound was recorded in course 000° with four different speeds, also was recorded with four different ranges and finally was recorded with four different directions degree. Every range and direction was measured in relation to own ship with course 000°. Now the data sets contain ($6\times 3\times 4 = 72$) audio files from 6 different types of simulated ships. Each file was recorded using mono-format with same microphone, same sound card, and has approximately durations of 9 seconds long. Each file was sampled at 44100 Hz; 16-bit quantization level was used. Next the data were segmented into approximately 23 ms frame length, overlapped by 50% of this frame. A Hamming was then applied to each frame.

Experiments & Results

Experiment description

The aim of this experiment is evaluating the effect of varying the number of coefficients for MFCC, PLPCC and RASTA-PLPCC on the total percentage of correctly classified targets (identification rate).

The experiment is organized as follows. First, split each recording file into intervals of 3 seconds for train and 3 seconds from the remaining file length for test, i.e.(3 train 3 test). Then, run the experiment using 8, 12, 16, 20, 24 and 26 coefficients for MFCC, PLPCC and RASTA-PLPCC

discarding 0th order cepstral coefficient. Finally, observe the identification rate for each extraction method while varying the number of coefficients used.

To ensure results, repeat the experiment using various intervals in seconds for the train and the test signals, e.g. 3 train 6 test, 6 train 3 test, 8 train 1 test, etc. as shown in figures 1, 2 and 3.

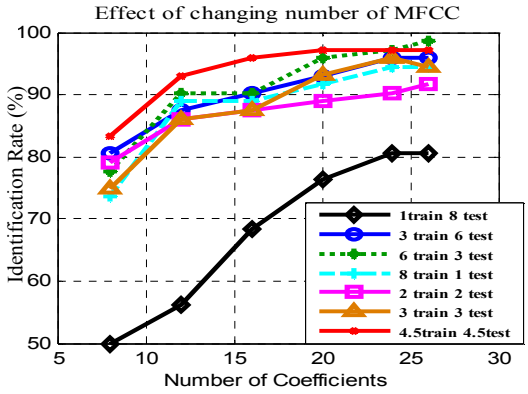

Figure 1: MFCC Identification Rate %

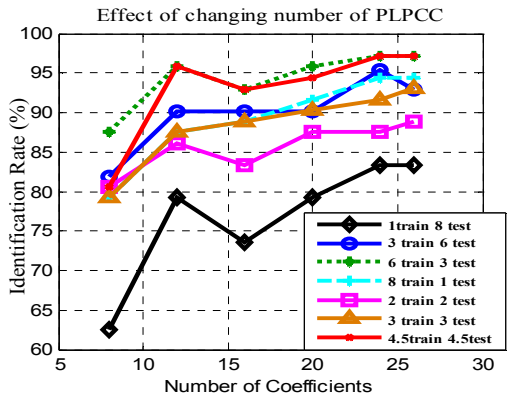

Figure 2: PLPCC Identification Rate %

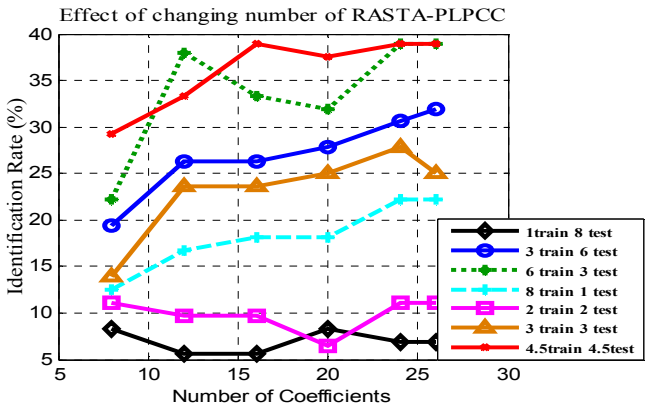

Figure 3: RASTA- PLPCC Identification Rate %

Results and Discussion

Figures 1, 2, and 3 show that for each feature type maximum identification rate is obtained at 24 coefficients and at 26 coefficients. Also, the identification rate at 24 coefficients differs slightly from that obtained at 26 coefficients. Higher is the number of coefficients, more is the time consumed within the identification process. Hence, the best choice is 24.

Figure 4 compare the identification rate for the three types of features while varying the number of coefficients used, either the train signal or the test one has a duration of 3 seconds. The experiment uses clear signals within the train and the test phases. Either MFCC or PLPCC yields to significantly higher identification rate than RASTA-PLPCC. Hence, MFCC and PLPCC seem to be better candidates. In addition, either PLPCC or RASTA-PLPCC needs computational time 2.3 times as much as MFCC does.

Conclusions

This paper presents three types of features, MFCC, PLPCC and RASTA-PLPCC to describe the passive sonar signal. Each feature type is used within the GMM to investigate the performance of the recognition process while varying the number of coefficients for each type. The processing time for extracting 26 coefficients is higher than that for 24 coefficients, whereas using 24 coefficients the identification rate differs slightly from that obtained using 26 coefficients. Hence, the best performance is obtained using 24 coefficients excluding the 0th coefficient. Moreover, the recognition results show that RASTA-PLPCC does not represent the ideal approach for clear underwater audio signals, but either MFCC or PLPCC seems to be a better candidate. Also, MFCC needs much less computational time than either PLPCC or RASTA-PLPCC.

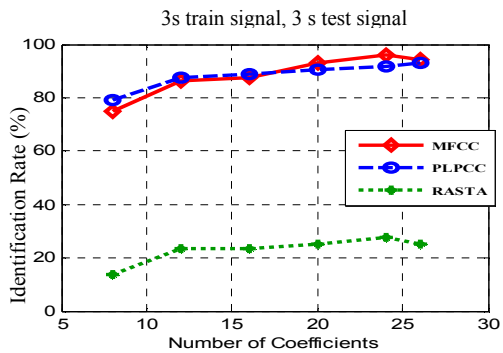

Figure 4: MFCC, PLPCC and RASTA- PLPCC Identification Rate %

References

[1] C. B Do, and S. Batzoglou, "What is the Expectation Maximization Algorithm?" Nature Biotechnology 26, pp 897 - 899, 2008.

[2] S. Davis, and P. Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences", IEEE Trans. Acoust., Speech, Signal Processing, Vol. 28(4), pp. 357–366, 1980.

[3] H. Hermansky, "Perceptual Linear Predictive (PLP) Analysis of Speech", Journal of the Acoustics Society of America, pp. 1738–1752, 1990.

[4] H. Hermansky, and N. Morgan, "RASTA Processing of Speech," IEEE Transactions on Speech & Audio Processing, Vol. 2, pp. 587–589, Oct. 1994.