Best
Available
Copy

(0)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

AD-A281 398

2. REPORT DATE
June 1994

3. REPORT TYPE AND DATES COVERED
Professional Paper

4. TITLE AND SUBTITLE
TRANSIENT SONAR SIGNAL CLASSIFICATION USING HIDDEN MARKOV MODEL AND NEURAL NET

5. FUNDING NUMBERS
PR: SUB8
PE: 0602314N
WU: DN308105

6. AUTHOR(S)
A. Kundu, G. C. Chen, and C. E. Persons

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Command, Control and Ocean Surveillance Center (NCCOSC)
RDT&E Division
San Diego, CA 92152–5001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Command, Control and Ocean Surveillance Center (NCCOSC)
RDT&E Division
San Diego, CA 92152–5001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In ocean surveillance, a number of different types of transient signals are observed. These sonar signals are waveforms in one dimension (1-D), and often display an evolutionary pattern over the time scale. The hidden Markov model (HMM) is well-suited to classification of such 1-D signals. Following this intuition, the application of HMM to sonar transient classification is proposed and discussed in this paper. Toward this goal, three different feature vectors based on autoregressive (AR) model, Fourier power spectrum, and wavelet transforms are considered in our work. The neural net (NN) classifier has been successfully used for sonar transient classification. The same set of features as mentioned above is then used with an NN classifier. Some concrete experimental results using "DARPA standard data set I" with HMM and NN classification schemes are presented. Finally, a combined NN/HMM classifier is proposed, and its performance is evaluated with respect to individual classifers.

Published in *International Conference on Acoustics, Speech, and Signal Processing*, pp. 325–328, April 1994.

14. SUBJECT TERMS
acoustic surveillance
antisubmarine warfare (ASW)

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAME AS REPORT

UNCLASSIFIED

21a NAME OF RESPONSIBLE INDIVIDUAL	21b TELEPHONE (include Area Code)	21c OFFICE SYMBOL
G. C. Chen	(619) 553-4924	Code 726

Accession For		
NTIS GRA&I	☑	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A-1	20	

NSN 7540-01-280-5500

Standard form 298 (BACK)

UNCLASSIFIED

ICASSP-94

VOLUME 2

S_2AUVN

SPEECH PROCESSING 2, AUDIO, UNDERWATER ACOUSTICS, VLSI & NEURAL NETWORKS

1994
IEEE INTERNATIONAL CONFERENCE
ON
ACOUSTICS, SPEECH AND
SIGNAL PROCESSING

APRIL 19-22, 1994
ADELAIDE CONVENTION CENTRE
ADELAIDE, SOUTH AUSTRALIA

IEEE

94CH3387-8

Sponsored by
THE INSTITUTE OF ELECTRICAL AND
ELECTRONICS ENGINEERS,
SIGNAL PROCESSING SOCIETY

TRANSIENT SONAR SIGNAL CLASSIFICATION USING HIDDEN MARKOV MODEL AND NEURAL NET

Amlan Kundu, George C. Chen, and Charles E. Persons

Naval Command, Control and Ocean Surveillance Center
RDT&E Division, Code 732
San Diego, CA 92152-5001, U.S.A.

ABSTRACT

In ocean surveillance, a number of different types of transient signals are observed. These sonar signals are waveforms in one dimension (1-D), and often display an evolutionary pattern over the time scale. The hidden Markov model (HMM) is well-suited to classification of such 1-D signals. Following this intuition, the application of HMM to sonar transient classification is proposed and discussed in this paper. Toward this goal, three different feature vectors based on autoregressive (AR) model, Fourier power spectrum, and wavelet transforms are considered in our work. The neural net (NN) classifier has been successfully used for sonar transient classification. The same set of features as mentioned above is then used with an NN classifier. Some concrete experimental results using "DARPA standard data set I" with HMM and NN classification schemes are presented. Finally, a combined NN/HMM classifier is proposed, and its performance is evaluated with respect to individual classifiers.

1. INTRODUCTION

The transient sonar signal classification problem is deemed difficult because of the short duration of the transients, wide intra-class variations and the effects of ambient ocean noise. The most common type of classifier used for this task is the neural net [1] though other classifiers have been studied [1, 4-5]. Also, it has been found that no single feature extraction technique can adequately capture all the feature information for all the ocean acoustic transients of interest. With this view in mind, we have experimented with the HMM classifier and three different feature vectors in this paper. The feature vector based on an AR model is a natural candidate with the HMM classifier. As the Fourier power spectrum is widely used by the NN community for their research, these features are also considered [1]. Finally, wavelet-transform-based features are considered. It is well-known that sonar transients are nonstationary signals. The wavelet transform can properly represent such signals. In particular, Daubechies type wavelets are considered in our work. These wavelets are finite duration filters and quite easy to implement [2]. It is our viewpoint that these three very different signal representations for feature extraction would reveal some of the latent characteristics of the signal for better classification.

Finally, we have studied the same set of features with a multi-layer perceptron neural net (MLP-NN) classifier with the express objective of finding out the complementary nature, if any, of these two classifiers - MLP-NN and HMM. We show in the current paper that a combined classifier using HMMs and MLP-NNs is likely to outperform the individual classifiers. Figure 1 gives the block diagram of our scheme.

2. FEATURE REPRESENTATION

We have three different feature representation schemes: one based on an autoregressive model, one based on Fourier power spectrum, and the other based on the wavelet transform. The AR coefficients are computed by the Burg algorithm. Due to scaling problem, the gain coefficient is not used.

2.1. Fourier Power Spectrum

From the given data segment, its FFT is computed. Before FFT computation, each data segment is windowed with a Kaiser-Bessel window function. The magnitude square of the FFT coefficients gives the Fourier power spectrum of the data.

2.2. Wavelet Transform

The Daubechies wavelets are a class of discrete orthonormal dyadic wavelets. An M order Daubechies wavelet [2] is given by M coefficients denoted by $C_j, j = 0, \ldots, M-1$. Then, the convolution of the signal with a FIR filter of length M $(C_j, j = 0, \ldots, M-1)$ gives the smooth component. On the other hand, the convolution of the signal with a FIR filter of length M and coefficients

$(-1)^{-j}C_{M-1-j}, j = 0, \ldots, M-1$, gives the detail component. After one pass of this algorithm, the smooth and the detail components are decimated by 2. The smooth components are then transformed again, and the procedure continues until we have only two smooth components left. The output, at this stage, is the wavelet transform of the original signal. The coefficients in Daubechies wavelets are obtained from orthonormality conditions, and "smoothness constraints". For an M order wavelet, these conditions and constraints lead to exactly M linear equations. Thus, M coefficients are uniquely determined [2].

2.3. Feature Selection

The feature representation schemes transform the original signal into feature space. Since some features may be more useful than others, only the important features should be selected for a compact representation of the signal for classification purpose. In our scheme, the signal is divided into a number of overlapping segments. All the AR coefficients are taken as the feature vector since relatively few AR coefficients are needed to represent a segment. For FFT power spectrum and the wavelet transform, the spectral and the transform coefficients with relatively higher magnitude are selected as features.

3. CLASSIFIER DESIGN

In our work, we have used two classifiers: HMM and multi-layer perceptron NN. Each signal template, i.e., exemplar, is divided into a sequence of partially overlapping segments. Each segment is then represented by one feature vector. The sequence of feature vectors, henceforth denoted as O, is used as one training/testing observation sequence for the HMM.

3.1. HMM Classifier

To solve our signal classification problem, we create one HMM for each class. The observation density in each state of the HMM is assumed to be multi-dimensional Gaussian. For a classifier of P classes, we denote the P models by $\lambda_p, p = 1, 2, \ldots, P$. When a signal O of unknown class is given, we calculate

$$p^* = \arg \max_{p} p(O, Q^* | \lambda_p)$$

and classify the signal as belonging to class p^*. Here, Q^* represents the optimal state sequence corresponding to O [3]. For a given λ, an efficient method to find $p(O, Q^* | \lambda)$ is the well known-Viterbi algorithm [3].

In creating the model for each class, we should guarantee that the parameters we obtain are the optimum for a given set of training samples. Since our decision rule is the state-optimized likelihood function, it requires that the estimated parameter $\hat{\lambda}$ be such that $p(O, Q^* | \hat{\lambda})$ is maximized over all possible λ for the training set. It is shown in [6] that the segmental K-means algorithm converges to the state-optimized likelihood function for a wide range of observation density functions, including the Gaussian density we have assumed.

In our works, a fully connected HMM topology is used. For the dataset used in our experiment, the fully connected HMM topology performs consistently better than the left-to-right HMM topology. However, there are sonar signals where the utility of left-to-right HMM topology has been demonstrated [4].

3.2. Multi-Layer Perceptron NN Classifier

Multi-layer perceptrons (MLP) are feed-forward nets with one or more layers of nodes between the input and output layers. The lowest layer is the input layer, which does not have any processing capability. The highest layer is the output layer and any layer between the input layer and output layer is called the hidden layer. It is the hidden layer that provides the MLP-NN classifier the ability to create highly nonlinear decision surfaces for better discriminative ability.

Generally, the multi-layer perceptrons are trained with the error back-propagation (EBP) algorithm [7] which is an iterative gradient algorithm designed to minimize the mean square error (MSE) between the desired output y_k^* and the actual output y_k. Sometimes, a momentum term is also included in the training procedure. In our scheme, there are 21 thirty-dimensional vectors in the sequence. These 630 features are used as training features for the NN. The NN is then designed with 630 input nodes, one hidden layer with 20 nodes, and trained using the back propagation algorithm and sigmoidal nonlinearity.

4. EXPERIMENTAL RESULTS

4.1. Signal Description

We have used the DARPA standard data set I for our experiments. This data set provides seven classes of signals to test our algorithm. A typical example, one from each class, is shown in Fig. 2. We denote these signal classes as:

Class A: Broadband 15-misc. pulse.

Class B: Two 4-misc. pulses, 27-misc. separation.

Class C: 3-kHz tonal, 10-misc. duration.

Class D: 3-kHz tonal, 100-misc. duration.

Class E: 150-Hz tonal, 1-sec. duration.

Class F: 250-Hz tonal, 8-sec. duration.

Class N: Ocean ambient noise.

We have created 45 templates, i.e., exemplars, for each class, of which 23 are used as training templates and 22 as test templates. Each signal template contains 1024 data points. The sampling rate for the signal is 24,576 Hz. For this sampling rate, 1024 data points are enough to capture the essential characteristics of all the transient types including the Class B type signal, which has the most time spread. This 1024 point signal template is divided into 21 frames, i.e., segments, of 256 data points with an overlap of 218 points (approximately 85%) between two successive frames. Once the feature vectors are computed from each frame, the signal template is represented by a feature vector sequence.

The training/testing sets include exemplars from four different SNR data sets. The SNR is computed as the ratio of the peak-signal-power to background-noise-power expressed in dB. The lowest SNR is 24 dB down with respect to the highest SNR. As a result, some very noisy exemplars are included in our experiments. We have tried a different number of states for HMM, from $N = 2$ to $N = 12$, and a different number of nodes, from 10 to 30, in the hidden layer of the NN. Only the best results are reported in the paper (Table 1). The results related to AR features are not reported as they are far inferior to those of other features.

4.2. Combined Classifier

Every feature/classifier combination has a somewhat different performance. A pertinent question is - can we combine the evidence of all the feature/classifier combinations to yield results that would be superior to any specific feature/classifier combination? Such a combined classifier would also be more robust. As a test, we have devised a classifier, henceforth called the majority classifier, that would take the output of each specific feature/classifier combination and assign the test exemplar the class with the majority votes only when the vote exceeds a threshold. Since we have six votes per test exemplar, we choose a threshold of 3 and 4. If the majority vote is below this threshold, that test exemplar is not classified. When the threshold is 4, only two test exemplars are misclassified, but 12 are not classified. When the threshold is 3, five test exemplars are misclassified, but only three are not classified.

Based on the detailed analysis of our experimental results, the following conclusions are in order:

(1) To a certain extent, the wavelet based features complement the FFT based features.

(2) To a certain extent, the HMM classifier complements the NN classifier.

(3) The combined classifier has the best result. Only a simple combination is described in the paper. Other possible combinations of HMM/NN classifiers should be explored. A hybrid HMM/NN classifier that combines the time normalization ability of the HMM and the superlative discriminative ability of the NN is currently being investigated.

Acknowledgment: The authors would like to thank Dr. Barbara Yoon of DARPA, Dr. Thomas McKenna of ONR and Code 535 at NRaD for the dataset.

REFERENCE

[1] J. Ghosh, L. M. 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.

[2] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," *Comm. in Pure and Applied Math*, Vol. 41, No. 7, pp. 909-996, 1988.

[3] Y. He and A. Kundu, "2-D Shape Classification Using HMM," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 13, 11, 1172-1184, 1991.

[4] J. P. Woodard, "Modeling and Classification of Natural Sounds by Product Code Hidden Markov Models," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. 40, No. 7, pp. 1833-1836, 1992.

[5] M. K. Shields, and C. W. Therrien, "A Hidden Markov Model Approach to the Classification of Acoustic Transients," *Proc. of ICASSP*, pp. 2731-2734, San Francisco, 1992.

[6] B. H. Juang and L. R. Rabiner, "The Segmental K-Means Algorithm for Estimating Parameters of Hidden Markov Models," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. 38, No. 9, September 1990, pp. 1639-1641.

[7] R. P. Lippmann, "An Introduction to Computing With Neural Nets," *ASSP Magazine*, pp. 4-21, April, 1987.

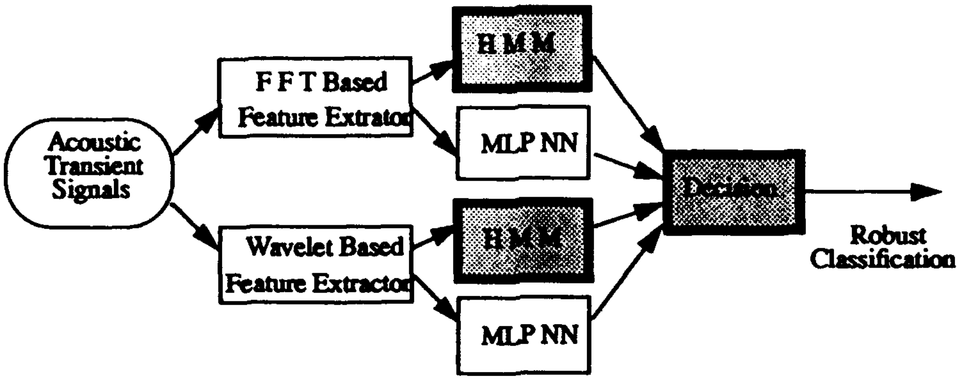

Figure 1. Block diagram representation of classification scheme.

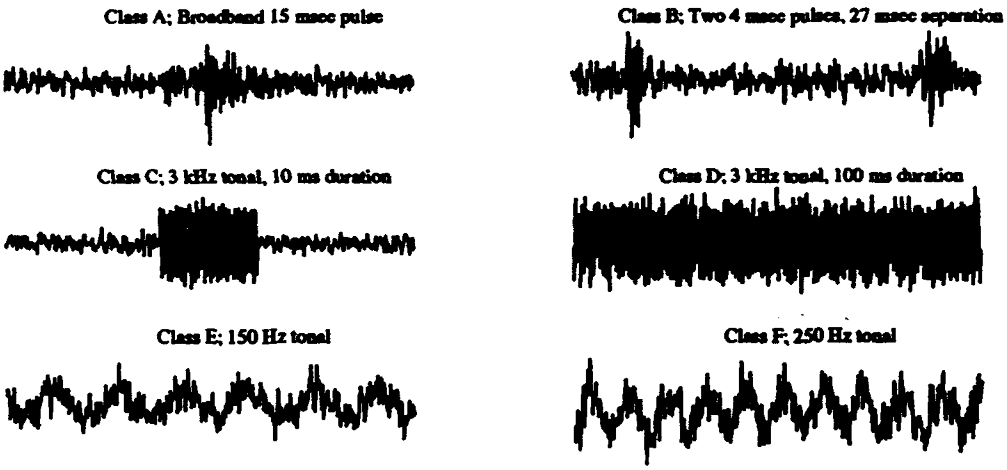

Figure 2. An example of different classes of signals used in our experiment.

	FFT	Wavelet (Daubechies 4)	Wavelet (Daubechies 20)	Combined
HMM	89.6%	91.5%	90.9%	-
NN	90.9%	90.9%	93.5%	-
Combined (Threshold=4)	-	-	-	*98.6%
Combined (Threshold=3)	-	-	-	*96.7%

Table1. Recognition performance of classifier/feature vector combination; * indicates that the "non-classified" templates are not included in computing the recognition performance.