

# A Study of Effects of Sonar Bandwidth for Underwater Target Classification

De Yao, Mahmood R. Azimi-Sadjadi, *Senior Member, IEEE*, Arta A. Jamshidi, and Gerald J. Dobeck

**Abstract**—The problem of classifying underwater targets is addressed in this paper. The proposed classification system consists of several subsystems including preprocessing, subband decomposition using wavelet packets, linear predictive coding, feature selection and neural network classifier. A multi-aspect fusion system is introduced to further improve the classification accuracy. The classification performance of the overall system is demonstrated and benchmarked on two different acoustic backscattered data sets with 40- and 80-kHz bandwidth. A comprehensive study is then carried out to compare the classification performance using these data sets in terms of the receiver operating curves, error locations, and generalization and robustness on a large set of noisy data. Additionally, the importance of different frequency bands for the wideband 80-kHz data is also investigated. For the wideband data, a subband fusion mechanism is introduced which offers very promising results.

**Index Terms**—Feature extraction, neural networks, target classification, wideband sonar.

## I. INTRODUCTION

THE PROBLEM of discriminating between underwater mines and similarly sized nontargets at ranges of several hundred meters remains a challenge. The objective is to discover what information is contained in the backscattered signals that might be exploited for target classification. Mine classification in shallow water is dominated by the needs to discriminate mines in reverberation and to separate them from competing clutter. The best way to accomplish this goal is through increasing sonar bandwidth as well as viewing the target over multiple aspects. Increasing the bandwidth gives greater resolution of the target and at the same time reduces reverberation by decreasing the effective scattering area of the bottom.

Designing a system for transmitting and receiving broadband signals is much more challenging than for narrowband sonar. The complexity increases substantially as the bandwidth increases, so careful attention should be paid to how much bandwidth is enough. For reducing reverberation and increasing definition of the target, the optimum resolution should be chosen such that most of the structure of the target can be separated, while the reverberation changes from a relatively smooth scattering function to distributed patches. Although

narrowband sonar relies on morphological clues in the signal to classify returns, wideband sonar offers the possibility of using spectral classification schemes at one or more scattering angles to estimate other target parameters such as face curvature, or to detect characteristic resonances and other spectral signatures. Wide bandwidth further allows the use of sparse arrays with their obvious advantages, and reduces the down-range size of a scattering cell. Additionally, environmental sensing can be done using broadband sonar to aid in real-time mission planning and dynamic adjustment of mine-hunting strategy.

A number of different schemes have been developed for classification of underwater targets from the acoustic backscattered signals. A review of the previous methods is provided in [1]. The method in this reference uses a wavelet packet-based classification scheme to discriminate mine-like and nonmine-like objects from the acoustic backscattered signals. The results on ten different noisy realizations with signal-to-reverberation ratio (SRR) of 12 dB indicated a very good classification performance on both single-aspect and multi-aspect decision cases. The system described in [2] was based on a new combination of simulated dolphin clicks, simulated auditory filters and artificial neural networks. Cylinders of various compositions were studied in the experiment. In this system, features from the echoes were extracted using a combination of a matched filter, envelope detection, a gammatone filterbank, time integration and Principal Component Analysis (PCA). The classification was performed by using a two-layer feedforward neural network with a modified SoftMax normalization. The results showed the best error rate as 5.9% when training samples are separated only by 2°. Huynh [3] presented a feature extraction method based on time-frequency analysis and Bienenstock, Cooper, and Munro (BCM) theory. Although the task in this reference was to classify the underwater mammals according to their emitted sounds, it presented a very interesting method for feature extraction. The input raw signals were first transformed using the Daubechies 4 wavelets. Training of a 10-BCM neuron using the transformed data was then performed. These neurons were fully connected and formed a network with lateral inhibition. After the training converged, the wavelet transformed data was transformed again using the BCM network. These, were then used as features to a subsequent classifier similar to the back propagation neural network. The results were compared with PCA method and showed a much better performance, i.e., 90% correct classification rate.

The goal of this paper is to develop an underwater target classification system and study its performance and in particular its robustness property for two different data sets with 40- and 80-kHz bandwidths. A three-aspect fusion algorithm is devised

Manuscript received June 27, 2000; revised February 5, 2002. This work was supported in part by the Office of Naval Research, the Biosonar Program under Contract N00014-99-1-0166.

D. Yao, M. R. Azimi-Sadjadi, and A. A. Jamshidi are with the Signal/Image Processing Laboratory, Department of Electrical Engineering, Colorado State University, Fort Collins, CO 80523 (e-mail: azimi@engr.colostate.edu).

G. J. Dobeck is with Signal/Image Processing, NSWC-Dahlgren Division, Coastal Systems Station, Panama City, FL 32407-7001.

Publisher Item Identifier S 0364-9059(02)05986-1.

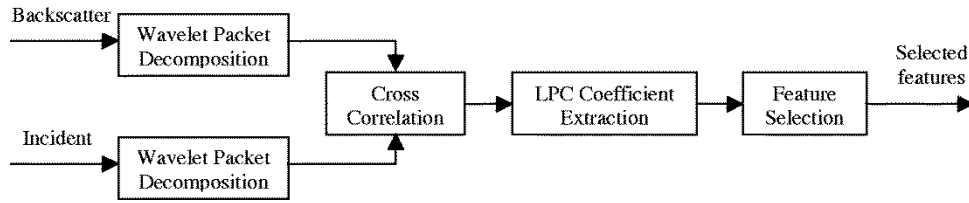


Fig. 1. Feature extraction and reduction processes.

to improve the classification rate and reduce the false alarm rate. Moreover, a study is conducted to determine the importance of different frequency bands for the wideband data set. Additionally, a new subband fusion scheme is proposed which provides excellent single-aspect target classification. This paper is organized as follows. A brief description of the classification system is presented in Section II. The single-aspect classification system and its results on the two data sets are presented in Section III. Two different multi-aspect fusion mechanisms are described and their results are analyzed and benchmarked in Section IV. For the 80-kHz data set, the importance of different frequency regions in the transmit signal and their effects on the classification performance are investigated in Section V. The subband fusion system is also introduced in this section. Finally, Section VI gives the conclusions and final remarks on the results of this paper.

## II. BRIEF DESCRIPTION OF THE FEATURE EXTRACTION/SELECTION SYSTEM

The steps involved in the front-end feature extraction and selection system are shown in Fig. 1. A wavelet packet-based scheme [4], [5] is used to decompose the frequency spectra of the backscattered signals into several subbands that contain useful target information. Wavelet packet decomposition provides an optimal multi-resolution decomposition of the signal spectrum in a manner very similar to the biological systems. The multi-resolution property allows for capturing fine details or subtleties in the signals the same way as zooming-in ability in the biological visual system in order to observe the small details in the detected objects. In WP decomposition, each subband extracts certain tonal features of the acoustic backscattered signals. To avoid phase distortion and at the same time ensure the orthogonality of the representation, Symlet 4 wavelet [4] is used in this study. Only those subbands that reside within the frequency range of the transmit signal, i.e., 20–60 kHz for the 40-kHz data and 30–110 kHz for the wideband 80-kHz data set, are selected. The transmit signal was also decomposed using the same WP tree structure.

In each selected subband, the cross-correlation between the transmit signal and the backscattered signal is performed as shown in Fig. 1. Then, the linear predictive coding (LPC) scheme [6] commonly used for speech coding and recognition applications is applied to the matched filtered result in each subband. For the 40-kHz data, a fourth-order linear autoregressive (AR) model,  $a_0y(n) = x(n) - \sum_{i=1}^4 a_iy(n-i)$  was employed while for the 80-kHz case a third-order AR model was used to represent the data. The model coefficients,  $a_i$ s and the gain  $a_0$  are used as features for classification.

Feature extraction is then followed by a feature reduction process that reduces the dimensionality of the feature space according to the discriminatory ability of each selected feature. To select an appropriate set of features, a criterion function can be used to provide the discriminatory power of the individual features. In this study, the Fisher discriminant function [7] was used to evaluate the distance between the two classes for each feature, i.e.

$$D_B(i, j) = \frac{|\mu_i - \mu_j|^2}{\sigma_i^2 + \sigma_j^2} \quad (1)$$

where  $\mu_i$  and  $\sigma_i^2$  represent the mean and variance of the features in class  $i$ , respectively. The features are sorted in a decreasing order of importance and the ones with higher discriminatory power were then selected to form the reduced feature vector. This process results in a reduced feature vector of 22 elements for both data sets. Note that the reduction of the wavelet features also removes the noise effects to some extent. For a more detailed description, the reader is referred to [1].

## III. CLASSIFICATION SYSTEM AND RESULTS

A two-layer back propagation neural network (BPNN) [8] with structure 22-42-2 (i.e., 22 inputs, 42 hidden layer nodes, and 2 output nodes) was used as a classifier. Classification decision was based upon discriminating mine-like from nonmine-like objects. The BPNN was trained using adaptive learning rate, momentum factor = 0.9, sum squared error goal = 10 and maximum number of epochs = 4000. Momentum decreases back-propagation's sensitivity to small details in the error surface in order to prevent the network from getting stuck in shallow minima, and the adaptive learning rate attempts to keep the learning step size as large as possible while keeping the learning stable. In this section, the single aspect classification results of the BPNN classifier are presented and analyzed. The performance comparison is made in terms of the receiver operating characteristic curves (ROC) and the error location plots for both data sets.

### A. Backscattered Data Sets

The data sets used in this study were collected by the Coastal Systems Station (CSS) in Panama City, FL. These data sets contain backscattered signals corresponding to six different objects—two mine targets namely a bullet-shape metallic object and a truncated-cone-shape plastic object; and four nontargets, namely a water-filled drum, an irregular shape limestone rock, a smooth granite rock, and a water-saturated wooden log. The transmit signal was a linear frequency

modulated (LFM) up-sweep with frequency range from 20 to 60 kHz for the first set and 30 to 110 kHz for the second one. Each object was insonified at aspect angles from 0 to 355 deg with 5-deg separation. This resulted in 72 aspect angles out of which the even-angles were used in the training data set while the odd-angle samples were used as the testing samples. As a result, for each object there are 36 patterns (at different aspect angles) in the training or testing data sets. The training data set contained the feature vectors of backscattered data with synthesized reverberation with  $SRR = 12$  dB that corresponds to nominal operating conditions. The procedure for generating synthesized reverberation involves convolving the transmit signal with a random sequence and scaling the resultant signal according to the specified  $SRR$  [1]. This reverberation signal is then added to the backscattered signal to generate one “noisy realization.” The process is repeated for every aspect angle multiple times in order to generate a statistically rich data set for determining the generalization ability of the classifier. The testing data in our study contained two sets of ten and fifty noisy realizations with  $SRR = 12$  dB for each aspect angle in the testing data in order to obtain statistically significant results. The results in the next section are obtained based on the ten-realizations testing set while the results for the robustness study in Section III-C use the second data set.

A five-level WP decomposition was then employed and 6 (for 40 kHz) or 12 (for 80 kHz) subbands that reside within the bandwidth of the transmit signal were selected. The transmit signal was also decomposed in the same way. Then, the cross-correlation of the subband signals of the backscattered and the corresponding subband signals of the transmit was performed as mentioned before. A fourth-order (or a third-order for the 80-kHz data) AR model was then fitted to the resultant data in each subband, and the model coefficients were extracted and used as features. This leads to a total of 30 (or 48) features out of which 22 features with high discriminatory power were selected according to the criterion in (1). These were then used for classification of the data.

### B. ROC Analysis and Classification Error of Different Classifiers

In this study, the data sets containing ten noisy realizations with  $SRR = 12$  dB was used to compare the performance of the classifier on the two data sets. The comparison is made in terms of their ROC curves and the error location plots versus aspect angle. The ROC curve is the plot of the probability of correct classification rate  $P_{cc}$  versus probability of false-alarm  $P_{fa}$ . Of particular importance is the behavior of the classifier at the “knee point” of this ROC that corresponds to a decision threshold leading to  $P_{cc} + P_{fa} = 1$ . The error location plot, on the other hand, gives the corresponding classification error locations and their frequencies versus aspect angles for each object. In this plot each ring corresponds to a particular object (i.e., objects 1–6) and each grid on the ring represents an aspect angle. In the counter-clockwise direction, the aspect angles start from 5 to 355 deg with 10-deg separation since there are 36 odd aspect angles for every object in the test set. The gray levels in each grid, which can vary from 0 to 10, represents the frequency of classification error at that particular aspect angle.

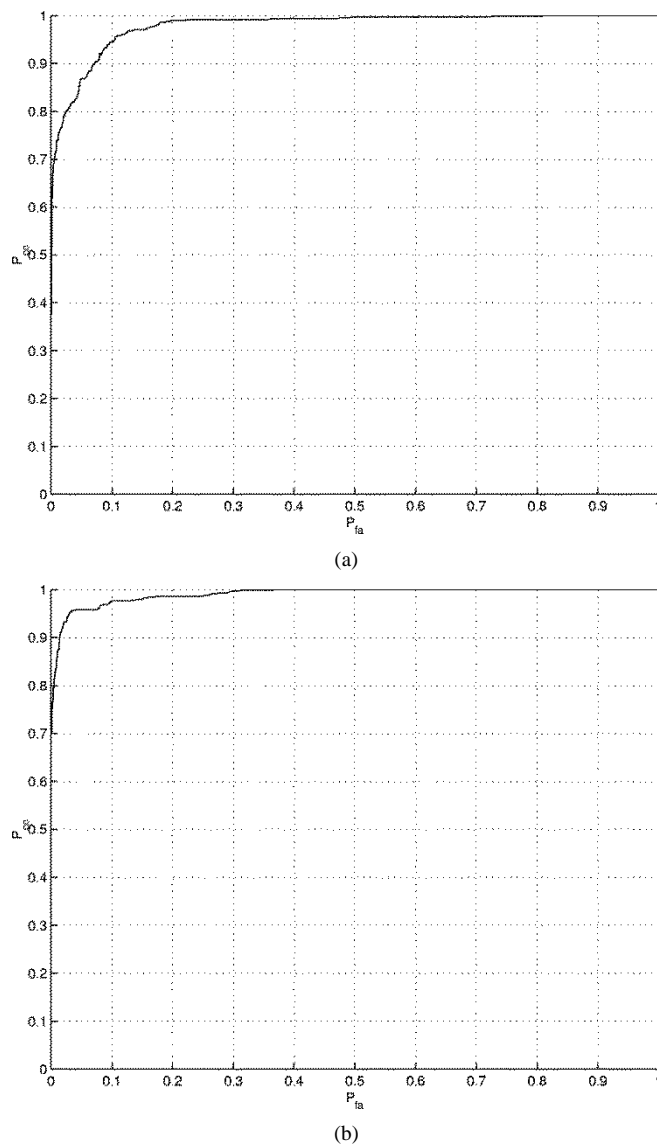


Fig. 2. ROC curves for the two data sets. (a) 40-kHz data set. (b) 80-kHz data set.

Fig. 2(a) and (b) shows the ROC curves for the 40- and 80-kHz data sets, respectively. As can be observed from these ROC curves, both cases led to excellent classification results based upon only the single-aspect information. The classifier trained and tested based upon the 40-kHz data set provided 91.94% correct classification ( $P_{cc}$ ) and 8.12% false alarm rates ( $P_{fa}$ ) at the knee of the ROC curve, while the one for the wide-band data gave slightly better performance, i.e.,  $P_{cc} = 95.83\%$  and  $P_{fa} = 4.24\%$  at the knee of the ROC curve. To study the classification performance for each individual target, the error location plots in Fig. 3(a) and (b) are generated for the 40- and 80-kHz data sets, respectively. The classification decision was made based on the threshold corresponding to the “knee” point of the ROC curves. As shown in Fig. 3(a), for the 40-kHz data, perfect classification results were obtained for objects 1, 5, and 6, i.e., cylindrical mine target, granite rock, and wooden log. Misclassifications occur very frequently at three aspect angles for object 3, i.e., 50-gallon water filled drum. All other errors occurred for objects 2 and 4, i.e., the second target and the

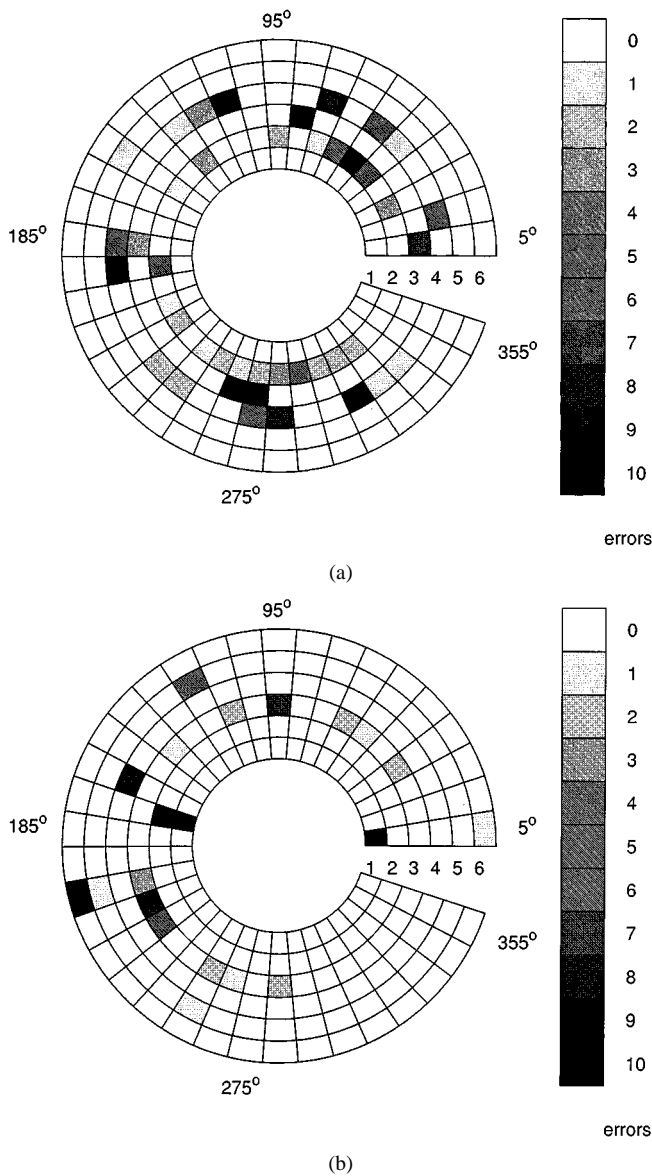


Fig. 3. Error location plots for the two data sets. (a) 40-kHz data set. (b) 80-kHz data set.

limestone rock, owing to their irregular shape and similarity of their features. For the 80-kHz data set, as evident from Fig. 3(b), the errors are mostly concentrated for target 1, the blunt-bullet-shape metallic object, object 3, the 50-gallon drum and object 6, the wooden log saturated in water. The differences between the results of the two data sets can be attributed to the fact that the target strength and signature varies as a function of beamwidth, frequency, grazing angle, etc.

### C. Robustness Analysis of Different Classifiers

To study the robustness and generalization of the classifiers to reverberation, we have studied the error rate statistics for a larger number of trials. For a given set of  $n$  input pattern vectors, the number of misclassifications  $m$  can be viewed as a random variable resulting in a randomly varying empirical error rate  $\epsilon_{\text{empirical}} = m/n$ . In [9], irrespective of the pattern source and the type of the classifier, it is shown that the random variable

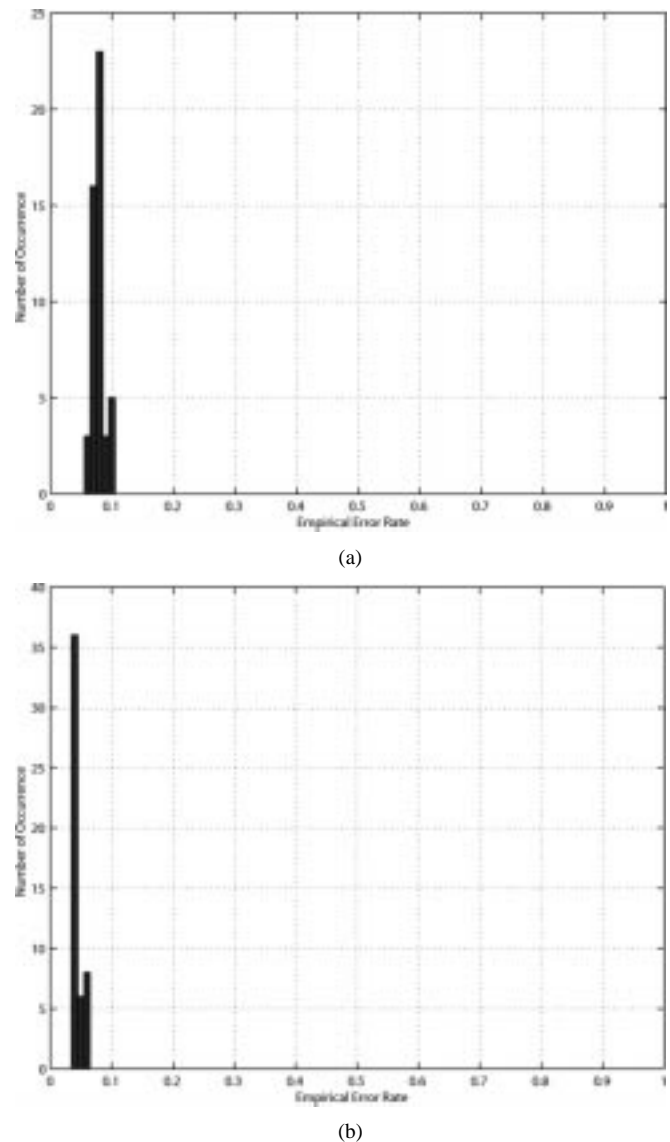


Fig. 4. Classification error statistics for fifty trials. (a) Classification error rate histogram for the 40-kHz data set. (b) Classification error rate histogram for the 80-kHz data set.

$\{m\}$  follows a binomial distribution. For a very large  $n$  this distribution approaches normal distribution. To estimate the error rate statistics of the classifiers, fifty different Monte Carlo trials were performed. For each trial, a different testing set with different reverberation realizations was used. As before, each set contained 216 backscattered signals corresponding to odd aspect angles and different synthesized reverberation sequences with SRR = 12 dB. Fig. 4(a) and (b) shows the histograms of classification error for the 40 and 80-kHz data sets, respectively. These results are obtained under the assumption that false positive and negative errors have the same weighting, i.e., the false alarm and misclassification errors are summed up to a total classification error. The classification decision is made based on the threshold at the “knee” point of the ROC curve. The classification error statistics (i.e., the mean,  $\mu$ , and the standard deviation  $\sigma$ ) are  $\mu = 0.0782$  and  $\sigma = 0.011$  for the 40-kHz data set and  $\mu = 0.0447$  and  $\sigma = 0.005$  for the 80-kHz data set, respectively. The results in Fig. 4(a) and (b) and the statistics clearly

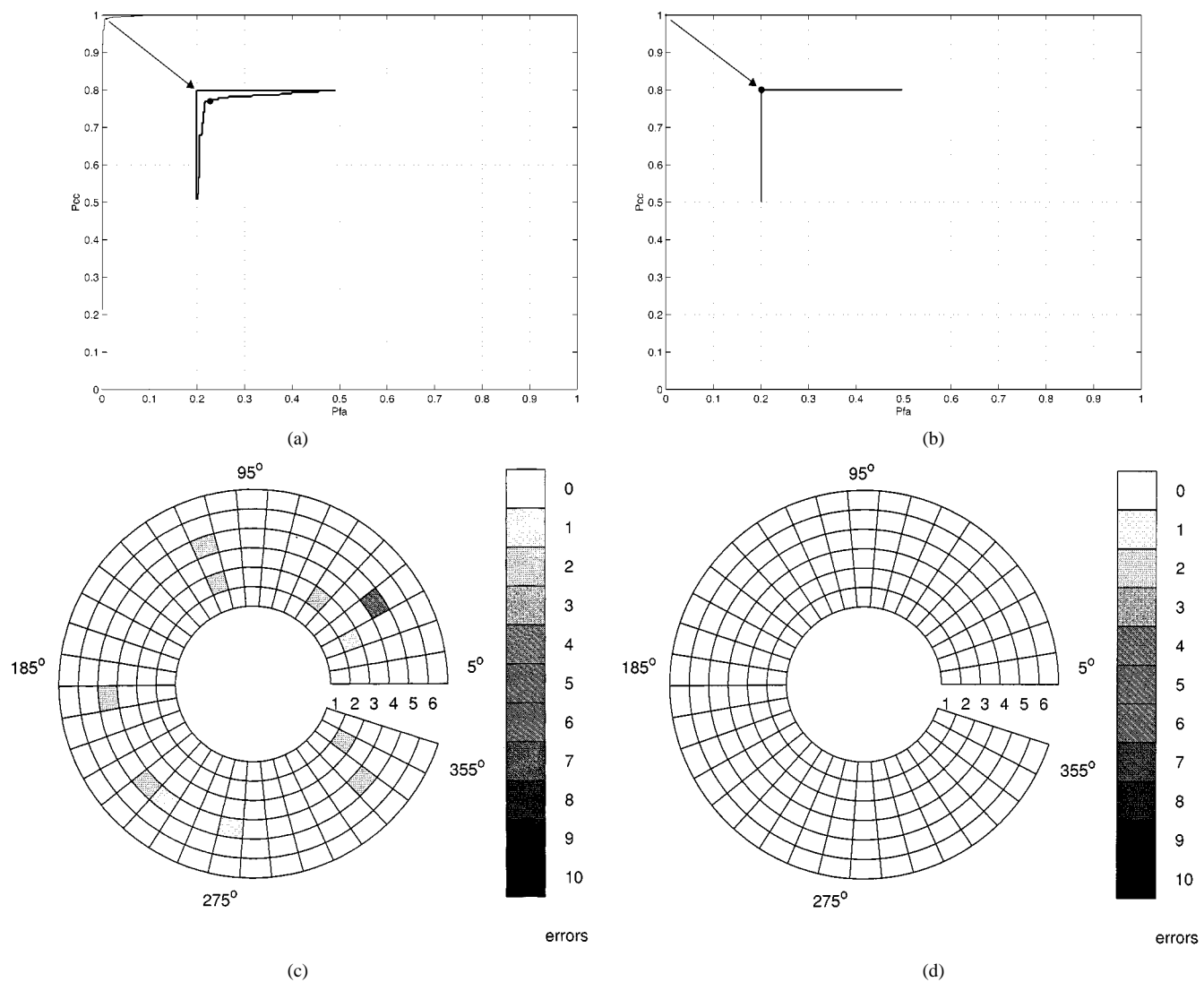


Fig. 5. ROC curves and error locations of the three-aspect fusion system for the two data sets. (a) ROC curve of the fusion system for the 40-kHz data set. (b) ROC curve of the fusion system for the 80-kHz data set. (c) Error location of the fusion system for the 40-kHz data set. (d) Error location of the fusion system for the 80-kHz data set.

show the improved robustness of the classifier to different reverberation realizations for the wideband data.

#### IV. MULTI-ASPECT FUSION SYSTEM AND RESULTS

The classification results of the two classifiers based on the single-aspect were fused together in order to further improve the overall classification performance. This is due to the fact that even though some target and nontarget cases may not be correctly classified at a specific aspect, it is very likely that they will be correctly classified at the neighboring aspects. Thus, the combined results of several aspects may lead to some improvements in the classification performance. In this study, similar to [1], [10]–[13], we have used sequences of three aspects separated by certain degrees. The reason is that in actual mine-hunting scenarios an object is declared as a potential mine if strong indications exist in 3–4 sonar pings. The three-aspect sequences are chosen so that overlap between them is allowed, e.g.,  $\{5, 35, 65\}$ ,  $\{35, 65, 95\}$ ,  $\dots$ ,  $\{355, 25, 55\}$  for  $30^\circ$  aspect separation interval. A study was also carried out to examine

the effects of varying the amount of aspect separation interval on the classification performance of the fusion system. The performance plot was generated when this separation interval was varied from  $10^\circ$  to  $40^\circ$  in  $10^\circ$  steps and the optimum case was selected. For each aspect, the two outputs of the classifier for target and nontarget are applied to the fusion system where the final decision is made based upon the classification results of the 3-aspect angles.

A two-layer back-propagation neural network was designed to perform the nonlinear three-aspect fusion. The network had six inputs corresponding to the classification outputs of the three aspects, four hidden layer nodes, and two output nodes for final decision. The single-aspect classification results of two different realizations based on the even aspect angle data set with different synthesized reverberation realizations were arranged in three-aspect sequence form, e.g.,  $\{0, 30, 60\}$ ,  $\dots$ ,  $\{350, 20, 50\}$  for  $30^\circ$  separation case. This data set (432 samples) was then used to train the neural network. The network reaches the sum squared error (SSE) goal of 10 within 200 epochs with adaptive learning rate and

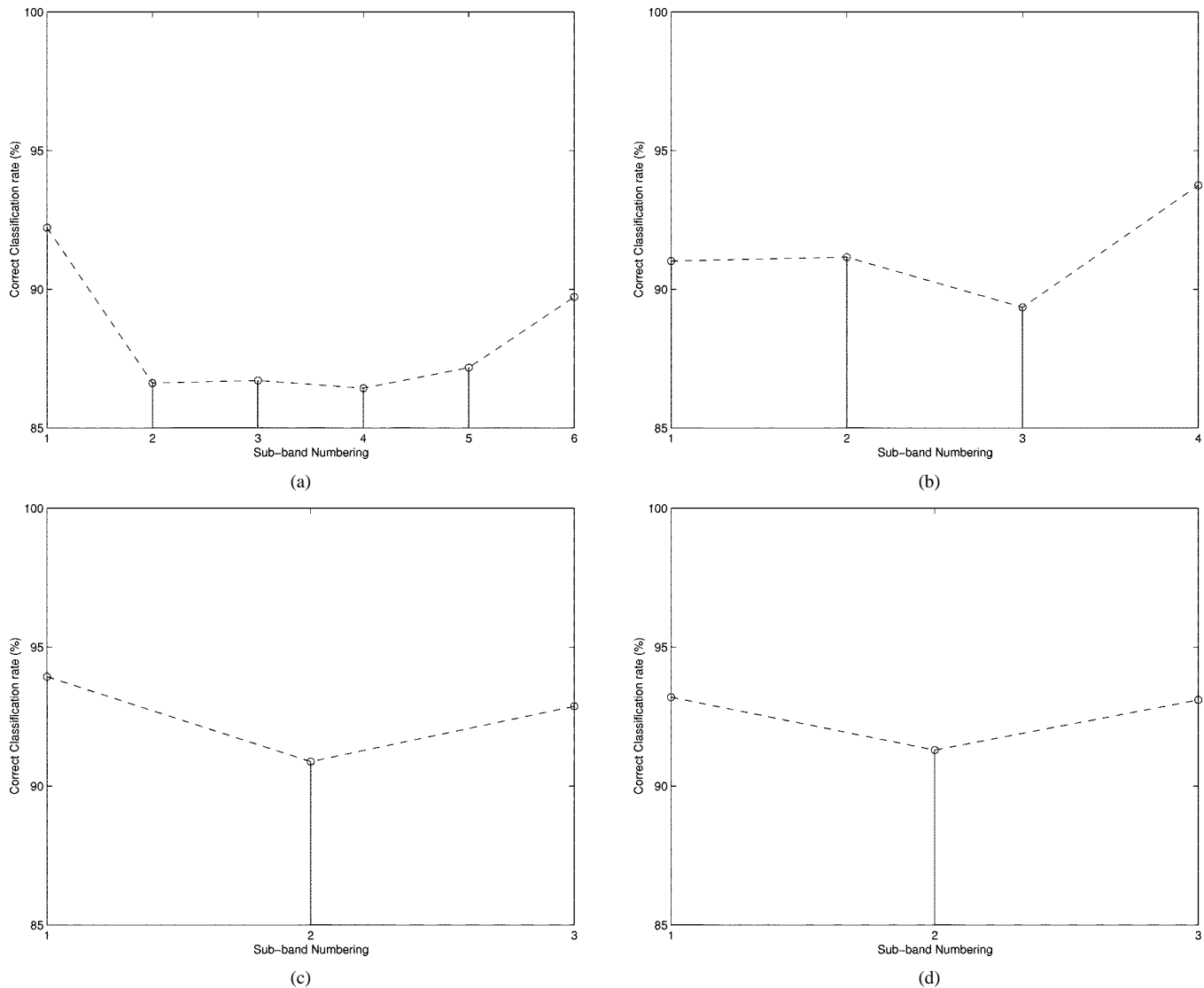


Fig. 6. Correct classification rates for different subband combination cases. (a) Case 1. (b) Case 2. (c) Case 3. (d) Case 4.

momentum of 0.9. The network was then tested on the testing data set that contains sequences of three odd aspects with the same aspect separation interval. For the 40-kHz data set, the best performance was obtained for  $40^\circ$  separation interval. As for the 80-kHz data set, the three-aspect fusion system led to perfect classification results at all the separation intervals tested. Fig. 5(a) and (b) shows the corresponding ROC curves of the three-aspect fusion results for this two-layer fusion network. For the 40-kHz data set,  $P_{cc}$  and  $P_{fa}$  at the knee of this ROC curve [Fig. 5(a)] are found to be 99.03%, 0.97%, respectively. Fig. 5(b) shows, once again, the perfect classification results for the wideband three-aspect fusion system.

Fig. 5(c) and (d) shows the corresponding error location plots of this fusion system for the 40- and 80-kHz data sets, respectively. In this case, the gray levels indicate errors of the system as a result of fusing the current aspect angle and its two previous aspect interval  $40^\circ$  for the 40-kHz data set, and  $30^\circ$  for the 80-kHz data set. Again the classification decision is made based on the threshold at the knee of the corresponding ROC curve. Comparing to the single aspect results in Fig. 3, it is evident that the multi-aspect fusion system significantly improved

the overall classification performance. For the 40 kHz, substantial improvements are obtained for objects 2 and 4. In addition, objects 1, 3, 5, and 6 are now correctly classified at all aspect angles in the testing data. As for the 80-kHz data set, substantial improvements are obtained for objects 1 and 3. In addition, objects 1, 4, 5, and 6 are now correctly classified at all aspect angles in the testing data, while only two errors occurred at  $5^\circ$  and  $145^\circ$  for object 3 and one error occurred at  $175^\circ$  for object 2.

In summary, all these results point to this important observation that the multi-aspect fusion indeed improves the classification performance. Furthermore, wideband insonification led to almost perfect target classification performance.

#### V. IMPORTANCE OF DIFFERENT FREQUENCY BANDS: 80-kHz DATA SET

The incident signal of the wideband data has a bandwidth that spans from 30 to 110 kHz. Since this range covers a wide bandwidth, the question that arises is that which frequency bands play dominant role in the classification decision for target versus nontarget discrimination. In order to find out the importance of

each frequency band, in terms of classification, the following study was carried out.

There were 12 subbands (out of 16) at the fifth level wavelet packet decomposition that reside in the frequency range of 30 to 110 kHz. Each subband occupies a bandwidth of 7.8125 kHz. In this study, we devised a scheme so that the subbands are used in different combinations. The features used here are the third-order LPC coefficients.

- *Case 1:* In this case, the subbands were grouped into groups of 2 subbands, i.e., {3, 4}; {5, 6}; {7, 8}; {9, 10}; {11, 12}; {13, 14}. The subbands collectively cover the frequency range from 23.4375 to 117.1875 kHz with each group covering a bandwidth of 15.625 kHz. All the LPC features corresponding to the two nodes in each combination were used for classification.
- *Case 2:* In this case, three subbands were combined together in a group, i.e., {3, 4, 5}; {6, 7, 8}; {9, 10, 11}; {12, 13, 14}. Each 3-subband group covers a bandwidth of 23.4375 kHz. Again, all the LPC features corresponding to each group of three nodes are used for classification.
- *Case 3:* In this case, four subbands were grouped together, i.e., {3, 4, 5, 6}; {7, 8, 9, 10}; {11, 12, 13, 14}. Each 4-subband group covers a bandwidth of 31.25 kHz. All the LPC features for these nodes are used for classification.
- *Case 4:* In this case, five subbands were grouped together. Since there are 12 subbands in total, overlapping had to be done. The combinations were: {3, 4, 5, 6, 7}; {8, 9, 10, 11, 12}; {10, 11, 12, 13, 14}. Each 5-subband group covers a frequency bandwidth of 39.0625 kHz. Again all the LPC features were used.

A two-layer BPNN classifier with different structures was used in each case listed above. The network structures were determined such that for a feature vector size of  $N$ , the hidden layer had  $(2 * N - 2)$  neurons while the output layer still had 2 neurons. The neuron activation functions were still the sigmoidal function and the training parameters were the same as those used for the other studies. The classification performance for these four cases are presented in the plots in Fig. 6 for different subband groups. Note that in this study since we are comparing the performance of different classifiers on the same data set all the classification decisions are made based upon hard-limiting decision, i.e., if  $o_1 > o_2$ , where  $o_1$  and  $o_2$  are the outputs of the BPNN, then the unknown pattern is classified as Target, otherwise is classified as Non-Target. As can clearly be seen, nodes {3, 4}; {5, 6} and {11, 12}; {13, 14}, which cover frequency bands from 23.4375 to 54.6875 kHz and 85.9375 to 117.1875 kHz, consistently provide much better classification than the nodes {7, 8} and {9, 10} which cover the frequency band from 54.6875 to 85.9375 kHz. This is evident by the appearance of a dip shown consistently in all the four cases. This study reveals the important observation that the useful information for classification may primarily be located in the sides of the bandwidth of the wideband data. This result may be exploited in the design of appropriate transmit signals for a specific wideband sonar target classification problem.

#### A. Subband Fusion Results: 80-kHz Data Set

In the study of the previous section, we saw that different frequency bands play different role in distinguishing targets from

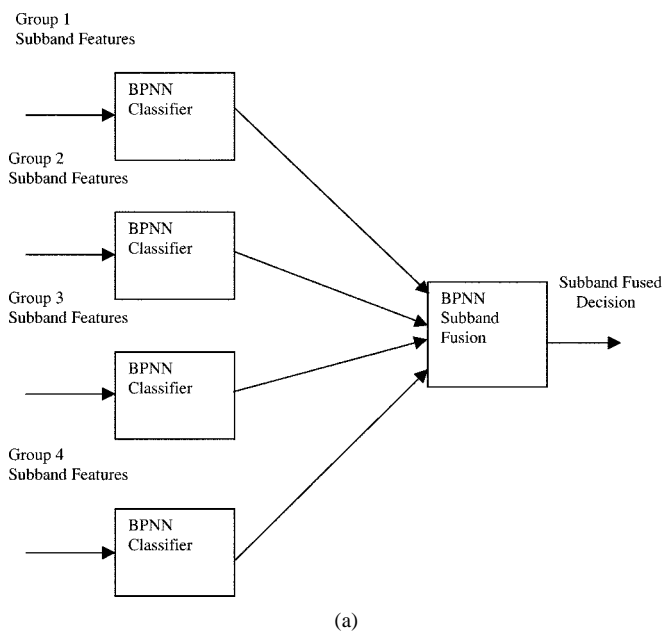


Fig. 7. Block diagram of the subband fusion system.

nontargets. In other words, the discrimination ability of each subband is dependent on its frequency range and the type of the target/nontarget. This property may be exploited to develop a “subband fusion” system in which the initial decisions based upon each individual subband are fused together to generate the final decision. This section demonstrates this idea on the wideband data set and compares the results of the subband fusion with those in Section III. In this study, the trained classifiers for Case 2 in the previous section were used. As explained before, for each group of three subbands, all the extracted LPC features in the associated subbands were combined together to form a 12-dimensional (four LPC coefficients per subband) feature vector. This gives four sets of feature vectors for the four groups. For each group, a two-layer BPNN with 12 inputs, 22 hidden layer neurons and two output neurons (12-22-2) was trained based upon the even aspect data associated with that group. The training was repeated for 36 trials with different randomly initialized weights in order to select the best trained network. This procedure was repeated for each group resulting in four trained networks. These networks were then tested on the testing data set that consisted of ten noisy (reverberation) realizations of the odd aspect data with SRR = 12 dB. The outputs of these four classifiers were fused together in another two-layer BPNN with 8 inputs, 20 hidden layer nodes, and 2 output nodes. This network, which was trained based upon the outputs of the preceding networks using the even aspect data, performs nonlinear fusion of the decisions. The entire system is shown in Fig. 7. This subband fusion system provided exceptionally superior classification results when compared to those in Section III. The ROC curve and the error location plot for this new fusion system are shown in Fig. 8(a) and (b). At the knee point of ROC, we have correct classification and false alarm rates of  $P_{cc} = 98\%$  and  $P_{fa} = 2\%$ , respectively. It must be emphasized that these results are obtained based upon only single-aspect decision. The error location plot also attests to this superior

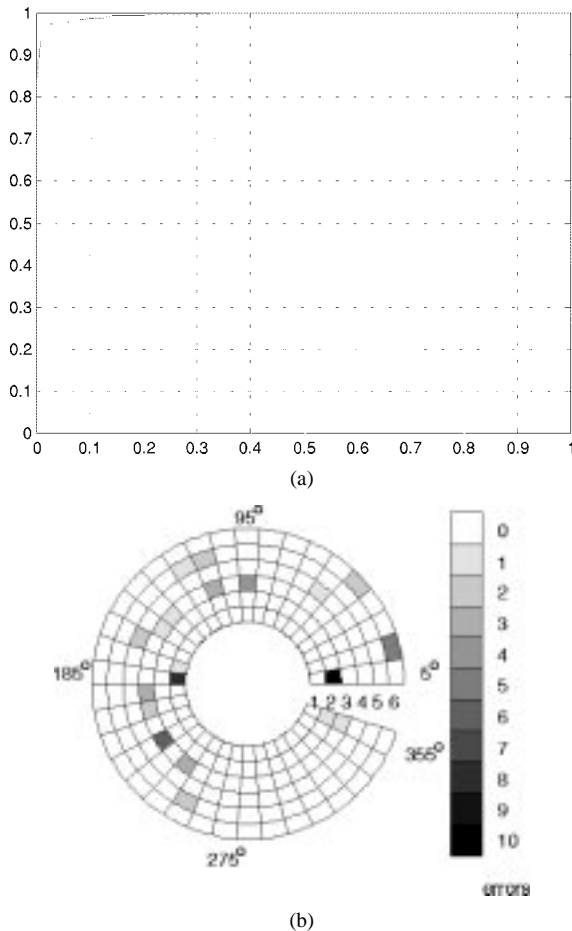


Fig. 8. Classification results of the subband fusion scheme for the 80-kHz data set. (a) ROC curve. (b) Error location plot.

performance. These results clearly show the promise of the subband fusion strategy.

## VI. CONCLUSION

The results presented in this paper indicate that the BPNN classifier performed very well for different conditions, in both single-aspect and multi-aspect cases. The wideband 80-kHz data provided much better discrimination than the 40-kHz data set. The single-aspect classification performance,  $P_{cc} = 95.83\%$ ,  $P_{fa} = 4.24\%$ , was considered to be very good for this data set. The three-aspect fusion system provided almost perfect classification performance by combining the results of three aspects. The generalization ability and robustness of the classifiers were also demonstrated on fifty noisy realizations with SRR = 12-dB conditions using Monte Carlo simulations. The study in Section V indicated that certain frequency regions in the wideband data play more important role in the classification than the others. It was observed that the middle frequency band of 55 to 86 kHz may be less important comparing to those in the sides, as far as target classification is concerned. This is an interesting point which may be exploited in the design of the transmit signals for a specific wideband

sonar system. Finally, we studied the idea of subband fusion for combining the decisions at different subbands for the 80-kHz data set. This study indicated that this new fusion paradigm provides exceptionally superior single-aspect classification results when compared to our previous results as well as to those of the others on this data set.

## ACKNOWLEDGMENT

The authors wish to thank NSWC, Coastal Systems Station, Panama City, FL, for the data and technical support.

## REFERENCES

- [1] M. R. Azimi-Sadjadi, D. Yao, Q. Huang, and G. J. Dobeck, "Underwater target classification using wavelet packets and neural networks," *IEEE Trans. Neural Networks*, vol. 11, pp. 784–794, May 2000.
- [2] L. N. Andersen, W. Au, J. Larsen, and L. K. Hansen, "Sonar discrimination of cylinders from different angles using neural networks," in *Proc. 1999 IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP '99)*, vol. 2, Phoenix, AZ, Mar. 1999, pp. 1121–1124.
- [3] Q. Q. Huynh, L. N. Cooper, N. Intrator, and H. Shouval, "Classification of underwater mammals using feature extraction based on time-frequency analysis and BCM theory," *IEEE Trans. Signal Processing*, vol. 46, May 1998.
- [4] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [5] S. G. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic, 1998.
- [6] L. Rabiner *et al.*, *Fundamentals of Speech Recognition*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [7] R. O. Duda, P. E. Hart, and D. Stork, *Pattern Classification*. New York: Wiley-Interscience, 2001.
- [8] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Englewood Cliffs, NJ: Prentice Hall, 1999.
- [9] J. Schurmann, *Pattern Classification, A Unified View of Statistical and Neural Approaches*. New York: Wiley-Interscience, 1996.
- [10] G. A. Carpenter and W. W. Streilein, "ARTMAP-FTR: A neural network for fusion target recognition, with application to sonar classification," in *Proc. SPIE, Detection and Remediation Technologies for Mines and Minelike Targets III*, vol. 3392, 1998, pp. 342–356.
- [11] C. F. Barnes, "Acoustic backscatter classification for mine detection using multiple fused aspects and novel database classification rules," in *Proc. SPIE, Detection and Remediation Technologies for Mines and Minelike Targets III*, vol. 3392, 1998, pp. 357–368.
- [12] D. Casasent and N. Kuljanyavivat, "Mine detection from multiple acoustic backscatter data," in *Proc. SPIE, Detection and Remediation Technologies for Mines and Minelike Targets III*, vol. 3392, 1998, pp. 370–381.
- [13] G. Okimoto and D. Lemonds, "Principal component analysis in the wavelet domain: New features for underwater object recognition," in *Proc. SPIE, Detection and Remediation Technologies for Mines and Minelike Targets IV*, vol. 3710, 1999, pp. 697–708.
- [14] J. A. Benediktsson, J. R. Sveinsson, O. K. Ersoy, and P. H. Swain, "Parallel consensual neural network," *IEEE Trans. Neural Network*, vol. 8, pp. 54–64, Jan. 1997.



**De Yao** received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, and the M.S. degree in electrical engineering from Colorado State University, Fort Collins, CO, in 1992, and 1999, respectively.

He worked as an Electronics Engineer with ShanDong Telecommunications Inc., ShanDong, China, from August 1992 to April 1996. From April 1996 until January 1998, he was a Hardware Engineer with Goldtron Telecommunications Private Limited, Singapore. He is currently with AirNet Communications Corporation, Melbourne, FL. His research interests include digital signal processing, classification and applications in communication systems.





**Mahmood R. Azimi-Sadjadi** (S'81–M'81–SM'89) received the M.S. and Ph.D. degrees from the Imperial College of Science and Technology, University of London, U.K., in 1978 and 1982, respectively, both in electrical engineering with specialization in digital signal/image processing.

He is currently a Full Professor at the Electrical and Computer Engineering Department of Colorado State University (CSU), Fort Collins. He is also the Director of the Digital Signal/Image Laboratory at CSU. His main areas of interest include digital signal

and image processing, target detection, classification and tracking, adaptive filtering and system identification, and neural networks. His research efforts in these areas resulted in over one hundred fifty journal and refereed conference publications. He is the co-author of the book *Digital Filtering in One and Two Dimensions*, (New York: Plenum Press, 1989).

Dr. Azimi-Sadjadi is the recipient of the 1999 ABELL Teaching Award, 1993 ASEE-Navy Senior Faculty Fellowship Award, 1991 CSU Dean's Council Award, and 1984 DOW Chemical Outstanding Young Faculty Award. He served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and is currently serving as an Associate Editor of IEEE TRANSACTIONS ON NEURAL NETWORKS.



**Arta A. Jamshidi** received the B.S. degree in electrical engineering, with Honors (*Summa Cum Laude*) from the University of Tehran, Tehran, Iran, in July 1998, and is currently working toward the M.S. degree at the Department of Electrical and Computer Engineering at Colorado State University, Fort Collins.

He is a Research Assistant at the Department of Electrical and Computer Engineering at Colorado State University. His research interests include signal/image processing, communications, and

pattern recognition.



**Gerald J. Dobeck** received the B.S. degree in physics from the University of Massachusetts, Amherst, in 1970, and the M.S. and Ph.D. degrees in electrical engineering from the University of South Florida, Tampa, in 1973 and 1976, respectively.

Since 1976, he has been employed at the NAVSEA Coastal Systems Station, Naval Surface Warfare Center, Dahlgren Division, Panama City, FL. His current research interests include automatic detection and classification of targets in cluttered environments from synthetic/real aperture sonar im-

agery, the echo structure of acoustic returns, underwater electro-optic imagery, and gradiometer/magnetometer signals. Over the past eight years, he is project leader for the development and application of processing technologies for: 1) automated mine detection and classification; and 2) data/sensor/algorithm fusion sponsored by the Office of Naval Research 6.2/6.3 Mine Countermeasures program. He has authored or coauthored more than 70 technical reports and papers.

Dr. Dobeck received the 1981 and 1996 Commanding Officer/Executive Director award for Science and Technology. In 2000, he received the National Defense Industrial Association Bronze Metal Award. He is a reviewer for IEEE, ASME, SPIE and the *Journal of Underwater Acoustics* and has been session chair at past IEEE and SPIE conferences.