# Comparison of sonar discrimination: Dolphin and an artificial neural network

Whitlow W. L. Au
*Marine Mammal Research Program, Hawaii Institute of Marine Biology, University of Hawaii, P.O. Box 1106, Kailua, Hawaii 96734*

(Received 22 June 1993; accepted for publication 19 January 1994)

The capability of an echolocating dolphin to discriminate differences in the wall thickness of cylinders (3.81 cm o.d. and 12.7 cm length) was determined by Au and Pawloski [J. Comp. Physiol. A **170**, 41–47 (1992)]. The dolphin was required to discriminate a standard target from comparison targets of differing wall thicknesses. Performance varied from 96% to 56% correct depending on the wall thickness of the comparison targets. The 75% correct threshold was determined to be wall thickness differences of −0.23 mm for comparison targets with thinner walls and +0.27 mm for comparison targets with thicker walls than the standard. The dolphin performance was unchanged in the presence of artificial broadband masking noise until the echo-energy-to-noise ratio fell below approximately 15 dB. A counterpropagation artificial neural network was used to examine broadband echo features from the same cylinders. Features of the echoes were determined by passing them through a filter bank of constant-$Q$ filters. Echo features of the standard and each comparison target were analyzed in pairs by a neural network having two output nodes. Twenty echoes per target were used in the training set and 30 additional echoes per target were used in the test set. For the noise free condition, the network performed at a comparable level to the dolphin for $Q$ values between 4 and 5. In the presence of noise, $Q$ values between 7 and 8 were needed before the network could perform at a comparable level to the dolphin for echo-energy-to-noise ratios of 10 and 15 dB. Discrimination cues seem to originate from a shift in the spectra of the thinner comparison targets to lower frequencies and to higher frequencies for the thicker comparison targets.

PACS numbers: 43.80.Jz, 43.80.Ka, 43.80.Lb

## INTRODUCTION

Echolocating dolphins have demonstrated a keen capability to discriminate, recognize and classify underwater targets (Nachtigall, 1980; Au, 1993). Comparable target discrimination capabilities have not been achieved with conventional sonar. Therefore, the use of artificial neural network to simulate the pattern recognition capability of the mammalian brain is attractive and may lead to the solution of problems that conventional AI and other methods have had only limited success. Gorman and Sejnowski (1988) successfully applied a back propagation network to a sonar problem classifying a metal cylinders and a cylindrically shaped rock. Roitblat *et al.* (1989) were the first to apply an artificial neural network to emulate a dolphin performing a sonar recognition task. The dolphin was involved in a delayed matching-to-sample task in which it was required to discriminate between a stainless-steel water-filled sphere, a solid aluminum cone and two water-filled plastic cylinders of the same length but different diameter. Echoes from the targets were first collected in a test pool using a simulated dolphin sonar signal as the incident signal. The echoes were then passed through a filter bank consisting of 20 constant interval bandpass filters, each having a bandwidth of 5 kHz. The amplitude of 20 bandpass filters were used as inputs to the network and the performance of the network was perfect with these echoes. A directional hydrophone was then used to collect echoes from the targets as the dolphin performed the sonar task. The counterpropagation network performed at a 97% correct level when selective "natural echoes" resulting from the dolphin's sonar emissions were used. A backpropagation network with 20 input units was also used with the "natural echoes" and performed at the same 97% correct level as the counterpropagation network.

Moore *et al.* (1991) extended the work of Roitblat *et al.* (1989) by developing an integrator gateway network to process consecutive rather than selective "natural echoes" with the dolphin performing the matching-to-sample task. The integrator gateway network consisted of a preprocessor unit and a backpropagation network. The gateway preprocessor unit determined the running average of 30 constant-frequency bandpass filters for consecutive echoes. The network was trained with six sets of ten successive echoes selected from the ends of randomly chosen echo trains. Two sets of echoes were chosen for each of three targets (large tube, sphere, and cone). The complete set of 1335 sequential echoes was presented to the trained network and the network was allowed to classify each echo train. The results were determined in term of "confidence ratio" defined as the ratio of the activation level of the correct classification unit versus the total activation level of the three classification units in the output layer. Overall, the dolphin's performance was better than the network, which achieved correct classification performance between 90% and 93%. However, the network tended to make cor-

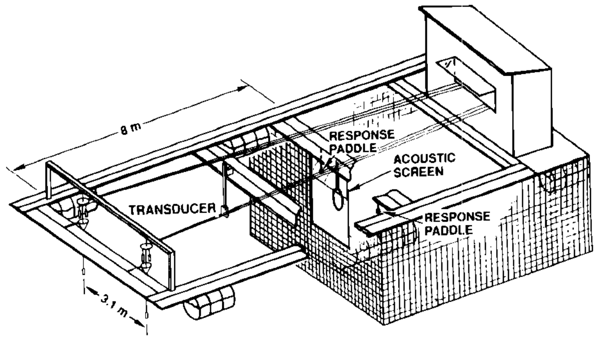


FIG. 1. Experimental geometry with the measurement transducer mounted on the dolphin's pen, 6 m from the targets.

rect classifications with fewer echoes than the dolphin.

In the study of Roitblat *et al.* (1990) and Moore *et al.* (1991), the dolphin's task was not difficult; the animal's average performance was 94.5% correct. The discrimination task in this study was considerably more difficult with the dolphin's accuracy varying from 96% to 53% correct. The dolphin was required to discriminate wall thickness differences of hollow cylinders by echolocation (Au and Pawloski, 1992). A standard cylinder of 6.35-mm wall thickness was compared with cylinders having wall thicknesses that differed from the standard by $\pm 0.2$, $\pm 0.3$, $\pm 0.4$, and $\pm 0.8$ mm. All cylinders had an o.d. of 37.85 mm, and a length of 12.7 cm. The standard and a comparison target, separated by an angle of 22°, were presented simultaneously at a range of 8 m and the dolphin was required to indicate the location of the standard target. The standard target was paired with a different comparison target for ten consecutive trials apiece. The experiment was conducted in both the free field and in the presence of broadband masking noise.

A counterpropagation artificial neural network was used to examine the broadband echo features from the same cylinders used in the dolphin experiment. Features of the echoes were determined by passing them through a bank of constant-$Q$ filters (the ratio of the center frequency and bandwidth of each filter is the same). Constant-$Q$ filtering was chosen because the dolphin's auditory system may be modeled as a bank of constant-$Q$ filters (Johnson, 1969). The objectives were to: (1) determine if a counterpropagation network could discriminate the target echoes using features from a constant-$Q$ filter bank, (2) compare the performance of the counterpropagation network with that of the dolphin, and (3) determine $Q$ values need by the network for comparable performance as the dolphin.

## I. PROCEDURE

Echoes from each of the targets used in the dolphin experiment of Au and Pawloski (1992) were measured with a monostatic echo collection system using a planar transducer to project and receive the acoustic signals. The transducer was mounted on the dolphin's pen located in Kaneohe Bay, Oahu, Hawaii, as shown in Fig. 1. Mounting the transducer on the dolphin's pen was done so that the target measurements would be made in the same environment and under similar conditions as for the dolphin. The most important effect of the environment consisted of pen and target motion caused by wind and waves. The pole supporting the transducer was aligned vertically and attached to the pen, with the transducer pointed directly at the target. Each target was suspended with a line in such a manner that its longitudinal axis was vertically aligned. With this geometry, the incident signal was nearly perpendicular to the longitudinal axis of the cylinders. The distance between the transducer and target was approximately 6 m. A simulated dolphin signal (Fig. 2) was projected at the targets and a Data Precision Data-6000 waveform analyzer operating at a sample rate 1 mHz was used to digitize the echoes with a resolution of 16 bits. Each echo consisted of 1024 digitized points which was stored on computer disk. Five disk files, each file consisting of ten consecutive echoes collected at a rate of five echoes per second, or 50 total echoes were collected for each target.

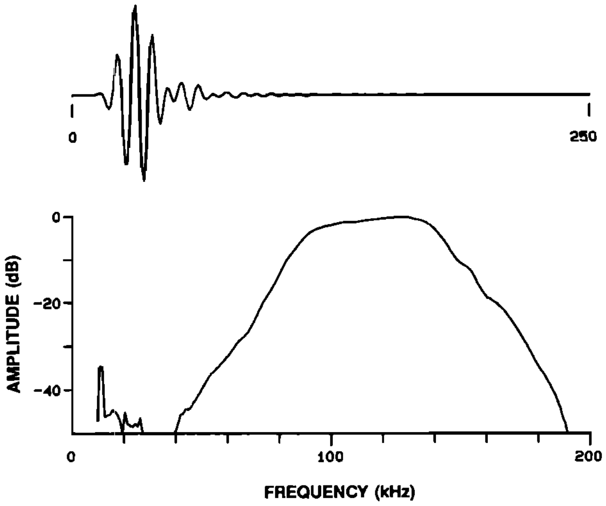


FIG. 2. Waveform and spectrum of the projected simulated dolphin echolocation signal.

Target features were determined by passing each echo through a bank of $N$ contiguous constant-$Q$ filters. The $Q$ of a filter is defined as

$$Q = f_0/\Delta f, \tag{1}$$

where $f_0$ is the center frequency and $\Delta f$ is the bandwidth. The center frequency of the $i$th filter is equal to $(f_{i+1}+f_i)/2$, and the bandwidth is equal to $f_{i+1}-f_i$, where $f_{i+1}$ is the upper frequency limit and $f_i$ is the lower frequency limit. From Eq. (1), the frequency boundaries of the $i$th constant-$Q$ filter can be expressed as

$$f_{i+1} = \frac{2Q+1}{2Q-1} f_i. \tag{2}$$

Let $f_U$ = the upper frequency of the $N$th filter and $f_L$ = the lower frequency of the 1st filter of a bank of constant-$Q$ filters, then from Eq. (1) the relationship between $f_U$ and $f_L$ can be expressed as

$$\frac{f_U}{f_L} = \left(\frac{2Q+1}{2Q-1}\right)^N. \quad (3)$$

For a specific $Q$, three parameters can be varied, $f_L$, $f_U$, and $N$. A frequency of 150 kHz was used for $f_U$ to coincide with the bottlenose dolphin upper frequency of hearing (Johnson, 1968). The lowest frequency was chosen so that, $f_L \geqslant 62$ kHz. For frequencies $<62$ kHz, the energy in an echo was at least 30 dB down from the peak. For a desired $Q$, $N$ was chosen so that $f_L$ was as close to 62 kHz as possible ($N$ was always 1 less than $Q$, for integer values of $Q<9$).

The energy output of each filter was determined by first calculating the fast Fourier transform (FFT) of each 1024-point echo, resulting in a discrete frequency spectrum defined at frequency intervals (bins) of 976.6 Hz. The boundaries of the constant-$Q$ filters were determined with Eqs. (2) and (3), and the energy of the $i$th filter was calculated using the following equation

$$E_i = \int_{f_{i-1}}^{f_i} |S(f)|^2 df, \quad (4)$$

where $E_i$ is the energy of the $i$th filter and $S(f)$ is the FFT of the echo. Values of the spectrum for constant-$Q$ filter boundaries between the discrete FFT frequency points were obtained with a cubic spline interpolation algorithm. Therefore, the features or exemplars of an echo consisted of the energy from each filter in the constant-$Q$ filter bank normalized to the output of the filter with the maximum energy.

The counterpropagation neural network first devised by Hecht-Nielson (1987), was simulated by the Neural Works Professional II Plus program from Neural Ware, Inc. The network consisted of an input layer of $N$ neurons or processing elements (PE's), a normalizing layer of $N+1$ elements, a Kohonen layer of $N$ elements and a Grossberg output layer with two elements each representing a specific category of targets. Echoes associated with the standard target were paired with echoes from each comparison target.

Target features determined by the constant-$Q$ filter bank were directed to the input layer of the network which buffered the input to the normalizing layer. The normalizing layer normalized the input vectors so that they laid on a hypersphere of constant radius before being transferred to the third, or Kohonen layer. Each unit in the normalizing layer also served a fan-out role to connect it with every PE in the Kohonen layer through a separate weight. The output of each Kohonen neuron was the weighted sum of its inputs. The Kohonen unit with the largest output was set to 1 and its weights adjusted by some fraction (defined by the learning rate parameter) of the difference between each weight and the value of the corresponding normalized input, making the weights more closely approximate the input. The outputs of the other Kohonen neurons were set to 0 and their weights were not adjusted. Therefore, the Kohonen layer served a "winner-take-all" function and basically classified the input vectors into similar grouping by adjusting its weights so that similar input vectors activated the same Kohonen neuron. The training involved a self-organizing algorithm that operated in an unsupervised mode. The outputs of the Kohonen layer were connected to a Grossberg outstar layer containing two PE's, one for each of the two input targets. The desired output of this layer was 1 for the PE representing the input target, and 0 for the other PE. This layer learned connections between its PE's and the PE's in the Kohonen layer in such a way as to map the set of winners in the Kohonen layer to the desired output. During training, the PE's in the Grossberg outstar used the Widrow–Hoff learning rule to adjust their weights, mapping the partitioned regions of the feature space described by the PE's in the Kohonen layer into the desired categories.

Features from twenty echoes for each target were used for the training set and features from the remaining 30 echoes of each target were used for the test set. The targets were considered in pairs, with the standard target paired off with different comparison targets. Training was performed in two stages, first in the Kohonen layer for a predetermined number of iterations, and then in the Grossberg outstar layer. The number of iterations required for training depended on the learning rate parameter and the similarity between the echoes of the standard target and the specific comparison target. In this study, the default number of iterations and learning rate parameters provided by Neural Ware for training the Kohonen layer were used. Three thousand iterations divided in three increments of 1000 iterations each were used in training the Kohonen layer, with the learning rate coefficients made progressively smaller in the 2nd and 3rd increments. Training of the Grossberg outstar layer started immediately upon completion of the 3000 iterations used to train the Kohonen layer. Training continued until the rms output error fell below 0.001. Typical training required between 3100 and 5000 iterations. The network's capability to discriminate the standard from each comparison target was determined for different values of $Q$ and $N$.

The performance of the network with noisy data was determined by first adding a different burst of Guassian random noise to each target echo. The noisy echo was then passed through the constant-$Q$ filter bank and the energy of each filtered was calculated. A noise burst was created by passing white noise through a cosine taper window (Otnes and Enochson, 1978) having a half-power width of 264 $\mu$s. A width of 264 $\mu$s was chosen to correspond to the dolphin's integration time of 264 $\mu$s determined by Au *et al.* (1988).

## II. RESULTS

### A. Neural network

The performance of the counterpropagation network for the free-field echoes and $Q$ values of 4 and 5 are shown in Fig. 3, along with the dolphin's performance. The number of filters, $N$, was equal to $Q-1$ because of the specific upper and lower frequency limits chosen for the filter bank. The network's performance for a $Q$ of 4 was not as good as the dolphin when comparing the standard target with the

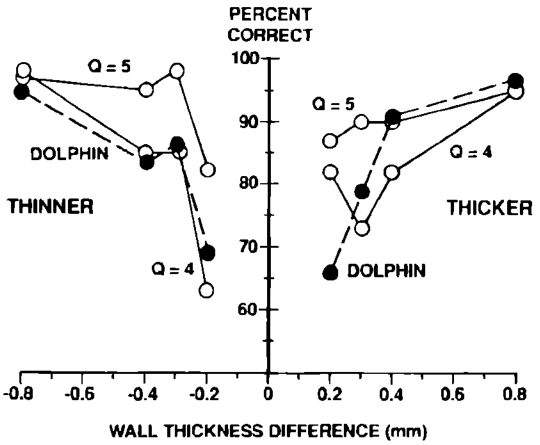


FIG. 3. Target discrimination results in the free-field for the neural network (open circles) using different constant-$Q$ values and for the dolphin (closed circles).

−0.3- and −0.2-mm comparison targets. However, for a $Q$ of 5 the network's performance was better than the dolphin for almost all of the thinner comparison targets. For the cases involving the thicker comparison targets, the results were slightly mixed. For a $Q$ of 5, the network performed better than the dolphin when comparing the standard with the 0.3- and 0.2-mm thicker targets, and slightly worse for the 0.4- and 0.8-mm thicker targets. However, for a $Q$ of 4, the network did poorer than the dolphin when comparing the standard target with the 0.3-, 0.4-, and 0.8-mm thicker comparison targets. These results indicate that, a constant-$Q$ filter bank consisting of four filters each having a $Q$ of 5 produced target features that allowed the network to perform as well as the dolphin in discriminating between the standard and comparison targets. The network results in Fig. 3 departed slightly from a monotonic decrease with decreasing wall thickness differences. These departures may have been caused by ping to ping variations in the echoes, which in turn may be attributed to motion of the pen and targets caused by wind and wave. Environmental noise was not a factor since the echo signal-to-noise ratio was very high.

The performance of the network after noise was added to the echoes is shown in Fig. 4 for signal-to-noise ratios of

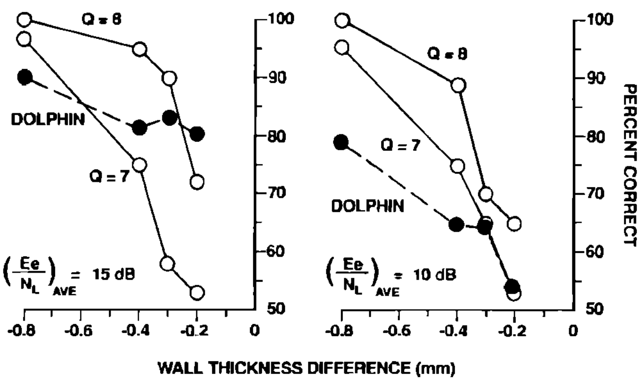


FIG. 4. Target discrimination results in noise for the neural network (open circles) and for the dolphin (closed circles). The averaged echo energy flux density to noise spectral density ratio is given on each graph.

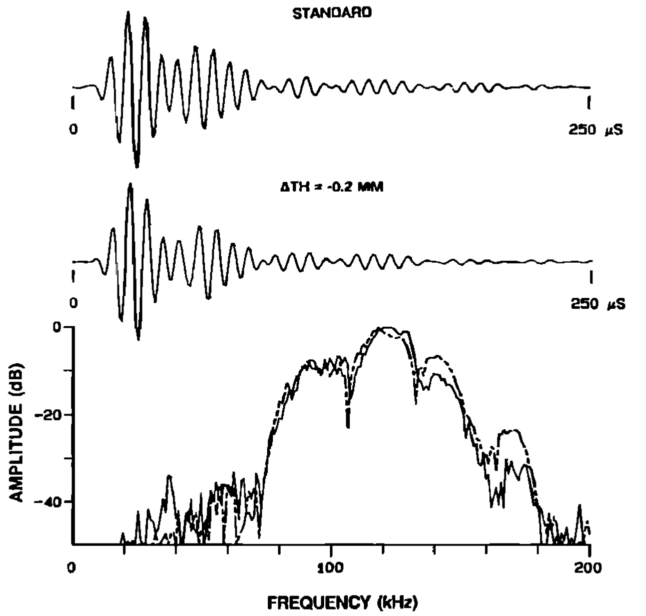


FIG. 5. Echoes from the standard cylinder and the comparison cylinder that was 0.2-mm thinner than the standard cylinder. The waveforms are shown in the top portion of the figure and the frequency spectra in the bottom portion. In the spectral plot the solid line is for the standard target and the dashed line is for the comparison target.

15 and 10 dB. Only the thinner targets were used in this study and the dolphin experiment (Au and Pawloski, 1992). For a $Q$ of 8 and a signal-to-noise ratio of 10 dB, the network performed better than the dolphin in discriminating the standard target from all the comparison targets. When the signal-to-noise ratio increased to 15 dB, the network performed better than the dolphin except for the −0.2-mm comparison target. For a $Q$ of 7 and a signal-to-noise ratio of 10 dB, the network generally performed better than the dolphin. When the signal-to-noise ratio increased to 15 dB the network was generally worse than the dolphin. The performance of the network did not improve much with the increased signal-to-noise, however, the dolphin's performance improved significantly when the signal-to-noise ratio increased from 10 to 15 dB.

### B. Spectral cues

Examples of echoes from the standard and the ±0.2-mm comparison targets are shown in Figs. 5 and 6. Small differences in the spectra for the comparison target (dashed line) compared to the standard target can be seen. The spectrum of the −0.2-mm comparison target seemed to be shifted slightly toward the left (toward lower frequencies) of the spectrum for the standard target. The shift in the spectrum of the comparison targets can be best seen by observing the nulls in the spectra for frequencies between 100 and 140 kHz. The spectrum of the +0.2-mm comparison target seemed to be shifted slightly toward the right (toward higher frequencies) of the spectrum for the standard target.

The shift in the position of the nulls and peaks in the spectra of the comparison targets can be explained by con-

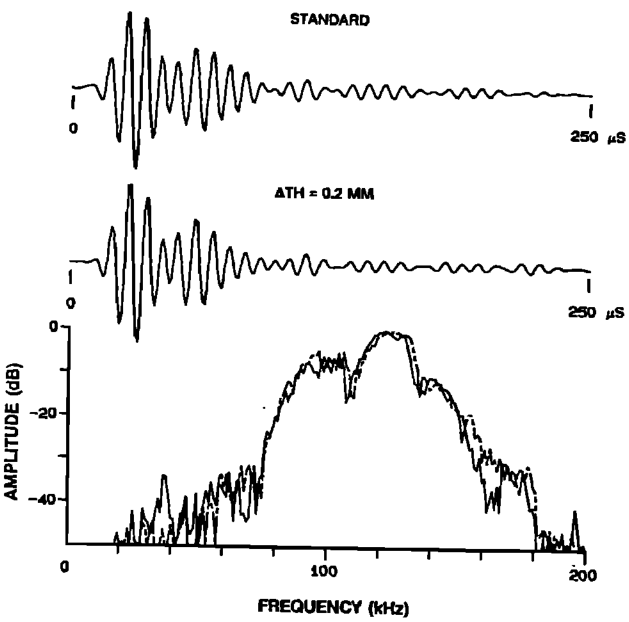


FIG. 6. Echoes from the standard cylinder and the comparison cylinder that was 0.2-mm thinner than the standard cylinder. The waveforms are shown in the top portion of the figure and the frequency spectra in the bottom portion. In the spectral plot the solid line is for the standard target and the dashed line is for the comparison target.

sidering the backscatter process depicted in Fig. 7. The first highlight is the result of reflection from the front wall of the cylinder and the second highlight is the result of the signal propagating through the target and reflecting off the interior back wall, propagating along the central path. Hammer and Au (1980) have demonstrated that the first two highlights of echoes from similar wall thickness cylinders will effectively dominate the shape of the spectrum. Therefore the spectrum can be approximated by considering a two highlight echo, which can be expressed as

$$e(t)=s(t)+k\cdot s(t-\tau), \tag{5}$$

where $s(t)$ is the first highlight, $s(t-\tau)$ is the second highlight multiplied by a constant $k$ which is less than 1, and $\tau$ is the time difference between the first and second highlights. The second highlight, $s(t-\tau)$ is exactly the same as $s(t)$ but occurs at a delay time $\tau$. The Fourier transform of Eq. (5) can be expressed as

$$E(f)=S(f)+e^{-j2\pi f\tau}S(f), \tag{6}$$

where $E(f)$ is the Fourier transform of $e(t)$ and $S(f)$ is the Fourier transform of $s(t)$. Both $E(f)$ and $S(f)$ are complex, having real and imaginary parts. The exponential term is the result of the shift theorem in Fourier analysis. Expanding the exponential term into cos and sin terms, the frequency spectrum of $e(t)$ can be expressed as

$$|\mathfrak{F}[s(t)+as(t-\tau)]|=\sqrt{1+a^2+2a\cos(2\pi f\tau)}\,|S(f)|, \tag{7}$$

where $\mathfrak{F}$ is the Fourier transform operator. The spectrum of the two highlight signal represented by Eq. (7) will have nulls when $\cos(2\pi f\tau)=-1$, which will occur at frequencies of

$$f_{\text{null}}=(2n-1)/2\tau, \tag{8}$$

where $n=1,2,3,...$ indicate the 1st, 2nd, 3rd, etc, null. Similarly, the spectrum will have peaks when $\cos(2\pi f\tau)=1$, which will occur at frequencies of

$$f_{\text{peak}}=(n-1)/\tau. \tag{9}$$

From Fig. 7, the time between the front wall reflection and the reflection off the inside back wall of the cylinder can be expressed as,

$$\tau=2\,\frac{\text{th}}{c_1}+2\,\frac{(\text{o.d.}-2\ \text{th})}{c_0}, \tag{10}$$

where th is the wall thickness and o.d. is the outer diameter of the cylinder, $c_1$ is the sound velocity in the aluminum wall and $c_0$ is the sound velocity of water. With $c_0=1530$ m/s and $c_1=5150$ m/s (Kinsler *et al.*, 1982), $\tau\approx 35.3\ \mu$s for the standard target. For targets used in this study, o.d.$\gg$th, so that the second term in Eq. (10) will be much larger than the first term because $c_1$ is about four times greater than $c_0$. Therefore, from the equation we can see that as wall thickness decreases, $\tau$ increases so that from Eqs. (8) and (9) the nulls and peaks will shift to lower frequencies causing the spectrum of the thinner targets to appear to be shifted toward lower frequencies. Conversely, as wall thickness increases, $\tau$ decreases so that the nulls and peaks shift to higher frequencies causing the spectrum of the thicker targets to also appear shifted toward higher frequencies.

The shift in the spectra of the comparison targets will cause the energy detected by some or all of the constant-$Q$ filters to be different for the comparison and standard targets. The average and standard deviation of the relative energy in each constant-$Q$ filter for a $Q$ of 5 and an $N$ of 4 is shown in Fig. 8 for all of the targets. Similarly, the relative energy in each constant-$Q$ filter for the situation depicted in Fig. 4 and a signal-to-noise ratio of 10 dB is shown in Fig. 9. Target echoes from both the training and test sets were used in obtaining the information shown in Figs. 8 and 9. The difference in the energy distribution between the standard and any of the comparison targets

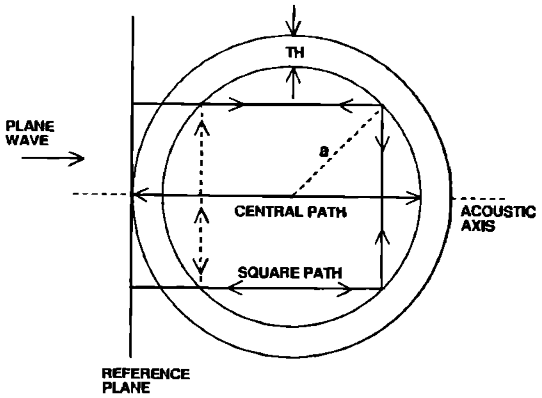


FIG. 7. Schematic showing the cross section view of a cylinder and the path of incident acoustic signal along with the central and square reflection paths within the target.

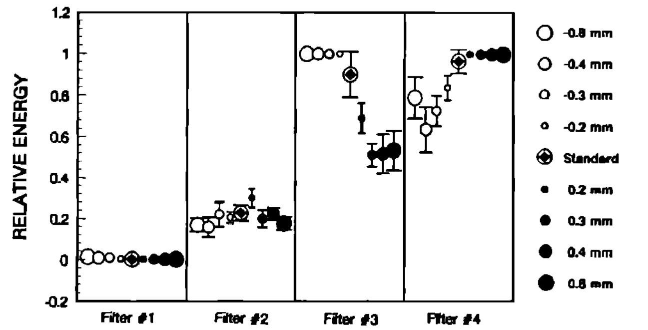


FIG. 8. The mean and standard deviation of the relative energy in each constant-$Q$ filter for the free-field condition. A filter bank with a $Q$ of 5 and 4 filters is represented.

can be observed in both figures. On the whole, the energy distributions of the comparison targets with wall thickness closest to the standard were most similar to that of the standard target.

### C. Euclidean distance measure

The wall thickness difference discrimination can be performed by other methods using frequency domain features. A simple and straight forward technique involves the use of a Euclidean distance measure to determine the similarity between the standard and comparison targets (Chestnut *et al.*, 1979). Let $\bar{E}_s(i)$ be the normalized energy in the $i$th filter averaged over all the standard target echoes in the training set, where $i=1,2,...,N$. Let $E_c(i)$ be the corresponding energy for a comparison target in the test set and $j$ be the index of the $j$th echo in the test set of $M$ echoes. The Euclidean distance measure $d_j$ of the $j$th comparison echo is defined as

$$d_j=\sqrt{\sum_{i=1}^{N} [\bar{E}_s(i)-E_c(i)]^2}. \tag{11}$$

For a specific comparison target, the smaller the value of $d_j$ the more similar the comparison target is to the standard target. Conversely, the larger $d_j$, the more dissimilar the comparison target is to the standard target. There will be $M$ values of $d_j$ associated with each comparison target. It should be noted that the Euclidean distance measure is also used in the Kohonen layer of the counterpropagation net-

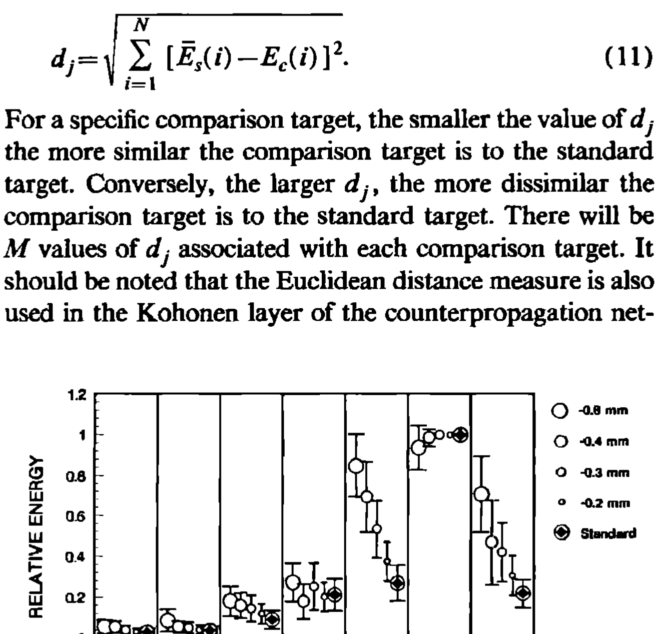


FIG. 9. The mean and standard deviation of the relative energy in each constant-$Q$ filter for a signal-to-noise ratio of 10 dB. A filter bank with a $Q$ of 8 and 7 filters is represented.

TABLE I. The mean and standard deviation of the Euclidean distance measure describing the similarity between the echoes for the standard target in the training or learn set and the echoes for the standard and comparison targets in the test set.

| | | THINNER COMPARISON TARGET | | | | No Noise Q = 5 N = 4 |
|---|---|---|---|---|---|---|
| | Standard (test) | -0.2 mm | -0.3 mm | -0.4 mm | -0.8 mm | |
| Standard (learn) | 0.16 ± 0.07 | 0.19 ± 0.07 | 0.29 ± 0.08 | 0.34 ± 0.09 | 0.22 ± 0.07 | $\bar{d}$ |

| | | THICKER COMPARISON TARGET | | | | No Noise Q = 5 N = 4 |
|---|---|---|---|---|---|---|
| | Standard (test) | 0.2 mm | 0.3 mm | 0.4 mm | 0.8 mm | |
| Standard (learn) | 0.16 ± 0.07 | 0.23 ± 0.04 | 0.40 ± 0.07 | 0.41 ± 0.08 | 0.42 ± 0.02 | $\bar{d}$ |

| | | THINNER COMPARISON TARGET | | | | SNR = 10 dB Q = 8 N = 7 |
|---|---|---|---|---|---|---|
| | Standard (test) | -0.2 mm | -0.3 mm | -0.4 mm | -0.8 mm | |
| Standard (learn) | 0.14 ± 0.07 | 0.21 ± 0.11 | 0.34 ± 0.20 | 0.49 ± 0.19 | 0.84 ± 0.19 | $\bar{d}$ |

work, to determine the difference between the stored weights of each Kohonen PE with the input features (Hecht-Nielsen, 1990).

The mean and standard deviation of the Euclidean distance measure between the standard and each comparison target in the test set is shown in Table I, for a $Q$ of 5 with no masking noise, and a $Q$ of 8 with a 10-dB signal-to-noise ratio. The results indicate that as the wall thickness difference between the standard and comparison targets increased the Euclidean distance measure also increased, except for the $-0.8$-mm thinner target. The standard target can be differentiated from the comparison targets provided a good threshold value for $\bar{d}$ is chosen.

## III. DISCUSSION AND CONCLUSIONS

The results indicate that in this particular study a counterpropagation network can discriminate target echoes as well or better than a dolphin, by using spectral cues obtained by processing the echoes with a bank of constant-$Q$ filters. We will now compare the $Q$ values needed by the network with that of the dolphin. Johnson (1969) measured the critical ratio of *Tursiops truncatus* at different frequencies between 5 and 100 kHz. Au and Moore (1990) augmented the critical ratio data of Johnson by measuring the critical ratio of *Tursiops* for frequencies up to 140 kHz. They also measure the critical bandwidth of the dolphin at 30, 60, and 120 kHz. The critical ratio and critical bandwidth measurements represent different estimates of the bandwidth of a mammalian auditory system. The results of both studies are shown in Fig. 10, with the data fitted to constant-$Q$ curves having the least-mean-square error. The critical ratio data was fitted with the $Q=12.5$ curve and the critical bandwidth data was fitted with the $Q=2.2$ curve.

The $Q$ values required by the counterpropagation network in order to perform as well as the dolphin are larger than the $Q$ represented by the critical bandwidth estimate but smaller than the $Q$ represented by the critical ratio estimate. The critical bandwidth measure may be more appropriate in an auditory discrimination task. The intelligibility of speech sounds by humans have been found to

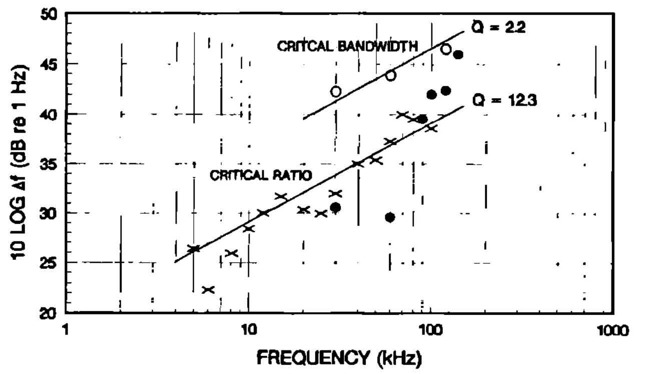


FIG. 10. Values of critical ratio (Johnson, 1969; Au and Moore, 1990) and critical bandwidth (Au and Moore, 1989) for *Tursiops truncatus*. The data are fitted to constant-$Q$ curves having the least-mean-square error.

be related to critical bandwidth (Celmer and Bienvenue, 1984; Bienvenue and Celmer, 1985). Speech signals with bandwidth less than a critical band were not as intelligible as speech signals with bandwidth equal to or greater than a critical band. If critical bandwidth is a more accurate estimate of the dolphin's auditory filter while discriminating sonar echoes, then the neural network required filters with a higher $Q$ than the dolphin, suggesting that the dolphin may be able to process the echoes with broader filters. However, this line of reasoning is highly speculative and what is required is a knowledge of the shape of the dolphin's auditory filter. Nevertheless, it is reassuring that the $Q$ values required by the network are within the two filter bandwidth estimates.

The network required more filters with higher $Q$ values when discriminating echoes in the presence of masking noise. The $Q$ increased from five in the noise-free case to eight in the noise case with the number of filters increasing from four to seven. This was not the case for the dolphin. Since the dolphin emitted signals that were on the average only 10-dB higher in amplitude when masking noise was used (Au and Pawloski, 1992), the difference in the echo levels with and without masking noise, should not be sufficiently great to change the dolphin's filter bandwidth. The echo level for a target having a −20-dB target strength located 8 m away will be approximately 57 dB below the level of the emitted signal. Weber (1977) and Lutfi and Patterson (1984) found that filter width in humans progressively broadens with signal level when the signal amplitude was above a relatively high level. Below this level, the filter width is relatively independent of signal level. Since the echo levels for the dolphin were not very high it is reasonable to assume that the dolphin's filter bandwidth remained essentially the same when masking noise was present and absent.

The fact that network required more filters with higher $Q$'s when discrimination targets in noise coupled with the likelihood that the dolphin's filters did not change in the presence of masking noise, suggest that the dolphin may be using both time domain and spectral cues. Although the target discrimination could be achieved by using only spectral cues originating from the shift in the spectra of the comparison targets, it seems reasonable that the dolphin would utilize whatever available cues and the combination of temporal and spectral information would maximize the efficiency of the discrimination process. One way of obtaining temporal-spectral information for a neural network is to convolve the echo signals with each constant filter and determine the energy in the temporal output of each filter.

The small number of relatively broad filters ($N=4$ and $Q=5$) required by the counterpropagation network in this study is surprising in light of the 20 constant bandwidth (5 kHz) filters used by Roitblat *et al.* (1989) and 30 constant bandwidth (3.9 kHz) filters used by Moore *et al.* (1991). Furthermore, the discrimination task was relatively easy in the study of Roitblat *et al.* (1989) and Moore *et al.* (1991) since the dolphin performed at the 94.5% correct level. For the targets used in this study, the dolphin's performance varied from 96% to 53% correct. The results obtained here certainly suggest that a smaller number of constant bandwidth filters could probably have been used in the other studies. However, there is one important difference between this and the other studies. In this study simulated, rather than actual dolphin echolocation signals which were the same from ping to ping were used to measure the target echoes. This is not the case when using "natural" echoes since the dolphin, being a biological entity and not a computer-controlled hardware system, will emit signals having slightly different frequency spectra from click to click. Furthermore, the echo waveform will be affected on any given ping by the specific animal-target orientation.

Caution must be taken in attempting to directly compare a dolphin's sonar performance in discriminating targets with that of a neural network. Although the counterpropagation network could perform as well as the dolphin, the filter bank had to have a higher $Q$ than the dolphin. Furthermore, the network had a relatively simple task of sorting 50 echo pairs in files that were already time windowed so that only target echoes were present. The dolphin had a more complex task of echolocating the targets, ignoring irrelevant echoes, determining the proper time window, remembering echo characteristics, as well as reporting to the experimenter. A better comparison would consist of a sonar system performing the discrimination task in real time, operating either simultaneously or in the same environment as the dolphin. Nevertheless, the results of this study indicate that the use of neural network technology may be a good technique in obtaining a better understanding of the dolphin sonar system and could provide insights as to the different cues available for target discrimination.

## ACKNOWLEDGMENTS

The author would like to express his appreciation to Peter Kaomea of NCCOSC RTD&E Division for suggesting a method to compare the energy distribution across the filters and to Dr. Herbert Roitblat, University of Hawaii, for his comments and suggestions. This study was sponsored in part by an Office of Naval Research Grant administered by Paul E. Nachtigall, contract number N0001492WX24139.

Au, W. W. L., Moore, P. W. B., and Pawloski, D. A. (1988). "Detection of Complex Echoes in Noise by an Echolocating Dolphin," J. Acoust. Soc. Am. **83**, 662–668.

Au, W. W. L., and Moore, P. W. (1990). "Critical Ratio and Critical Bandwidth for the Atlantic Bottlenose Dolphin," J. Acoust. Soc. Am. **88**, 1635–1638.

Au, W. W. L., and Pawloski, D. A. (1992). "Cylinder Wall Thickness Difference Discrimination by an Echolocating Atlantic Bottlenose Dolphin" J. Comp. Physiol. A **170**, 41–47.

Au, W. W. L. (1993). *The Sonar of Dolphins* (Springer-Verlag, New York).

Bienvenue, G. R., and Celmer, R. D. (1985). "Critical Bands in the Perception of Speech Signals. Part 2: SNHL Data," J. Acoust. Soc. Am. **78**, S41(A).

Celmer, R. D., and Bienvenue, G. R. (1984). "Critical Bands in the Perception of Speech Signals," J. Acoust. Soc. Am. **76**, S14(A).

Chestnut, P., Landsman, H., and Floyd, R. W. (1979). "A Sonar Target Recognition Experiment," J. Acoust. Soc. Am. **66**, 140–147.

Gorman, R. P., and Sejnowski, T. J. (1988). "Learned Classification of Sonar Targets using a Massively-parallel Network," Neural Networks **1**, 75–89.

Hammer, C. E. Jr., and Au, W. W. L. (1980). "Porpoise Echo-recognition: An Analysis of Controlling Target Characteristics," J. Acoust. Soc. Am. **68**, 1285–1293.

Hecht-Nielsen, R. (1987). "Counter-Propagation Networks," IEEE First Int. Conf. Neural Networks **II**, 19–32.

Hecht-Nielsen, R. (1990). *Neurocomputing* (Addison-Wesley, Reading, MA).

Johnson, C. S. (1968). "Masked Tonal Thresholds in the Bottlenose Porpoise," J. Acoust. Soc. Am. **44**, 965–967.

Johnson, C. S. (1969). "Sound Detection Thresholds in Marine Mammals," in *Marine Bio-Acoustics*, edited by W. N. Tavoga (Pergamon, New York), pp. 247–260.

Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (1982). *Fundamentals of Acoustics* (Wiley, New York).

Lutfi, R. A., and Patterson, R. D. (1984). "On the Growth of Masking Asymmetry with Stimulus Intensity," J. Acoust. Soc. Am. **76**, 739–745.

Moore, P. W. B., Roitblat, H. L., Penner, R. H., and Nachtigall, P. E. (1991). "Recognizing Successive Dolphin Echoes with an Integrator Gateway Network," Neural Networks **4**, 701–709.

Nachtigall, P. E. (1980). "Odontocete Echolocation Performance on Object Size Shape and Material," in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York), pp. 71–95.

Otnes, R. K., and Enochson, L. (1978). *Applied Time Series Analysis* (Wiley, New York), Vol. 1.

Roitblat, H. L., Penner, R. H., and Nachtigall, P. E. (1990). "Matching-to-Sample by an Echolocating Dolphin," J. Exp. Psychol. Anim. Behav. Proc. **16**, 85–95.

Roitblat, H. L., Moore, P. W. B., Nachtigall, P. E., Penner, R. H., and Au, W. W. L. (1989). "Natural Echolocation with an Artificial Neural Network," Int. J. Neural Networks **1**, 239–248.

Weber, D. L. (1977). "Growth of Masking and the Auditory Filter," J. Acoust. Soc. Am. **62**, 424–429.