DISCRIMINATION OF CYLINDERS WITH DIFFERENT WALL THICKNESSES USING NEURAL NETWORKS AND SIMULATED DOLPHIN SONAR SIGNALS

Lars Nonboe Andersen, Whitlow Au (*), Jan Larsen and Lars Kai Hansen

Department of Mathematical Modeling, Building 321
Technical University of Denmark, DK-2800 Lyngby, Denmark
Email: lna,jl,lkhansen@imm.dtu.dk
Web: http://eivind.imm.dtu.dk

(*) Marine Mammal Research Program, Hawaii Institute of Marine Biology
University of Hawaii, P.O. 1106, Kailua, Hawaii, 96734, USA

Abstract. This paper describes a method integrating neural networks into a system for recognizing underwater objects. The system is based on a combination of simulated dolphin sonar signals, simulated auditory filters and artificial neural networks. The system is tested on a cylinder wall thickness difference experiment and demonstrates high accuracy for small wall thickness differences. Results from the experiment are compared with results obtained by a false killer whale (*pseudorca crassidens*).

INTRODUCTION

Dolphins have demonstrated excellent capabilities of detecting, discriminating, recognizing and classifying underwater targets (see e.g. [3], [16]). Comparable target discrimination capabilities have not been achieved with conventional sonar. This has inspired research in new sonar systems based on biological knowledge, i.e. modeling dolphins' discrimination capabilities (see e.g., [15] and [21]).

The fact that the inner ear of the dolphin has many similarities with the human inner ear makes it tempting to use knowledge from simulations of the human auditory system when trying to model the dolphin sonar system. Furthermore, neural networks have proven to be very useful for pattern recognition tasks (see e.g., [21]) and are here applied for classification of features extracted from auditory preprocessing.

Based on earlier work [4], [14], [19] we will describe and present results from an experiment using a simulated dolphin signal, preprocessing based on auditory modeling and classification using a neural network. The aim is to discriminate between echoes recorded from hollow cylinders with different wall thicknesses.

Results from the experiment will be compared with results from a similar experiment involving a false killer whale (*pseudorca crassidens*) in which the whale was required to discriminate wall thickness differences of the same hollow cylinders used in our artificial experiment.

DATA COLLECTION

Echoes from ten hollow cylinders made of stainless steel were measured with an echo collection system using a planar broadband transducer to project and receive the acoustic signals. All data were collected in Kaneohe Bay, Oahu, Hawaii.

All cylinders had an outer diameter of 38.1 mm, and a height of 12.7 cm. One cylinder had a wall thickness of 6.35 mm and is denoted the standard cylinder. The other nine cylinders differed from the standard by 0 mm, ± 0.08 mm, ± 0.15 mm, ± 0.23 mm and ± 0.31 mm within an accuracy tolerance of 0.03 mm and are denoted comparison cylinders. A schematic of the standard and the comparison cylinders is shown in Fig. 1.

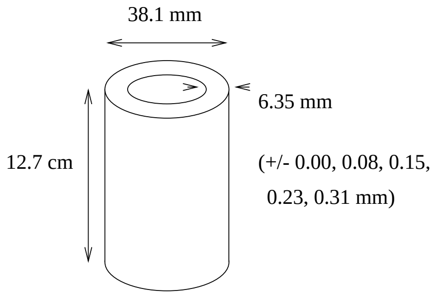

Figure 1: Schematic of the standard and comparison cylinders.

The transducer was mounted on a floating pen located in Kaneohe Bay. The pole supporting the transducer was aligned vertically and attached to the pen, with the transducer pointing directly towards the center of the suspended target. Each target was suspended with a monofilament line in such a manner 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 the target was approximately 5 m. An illustration of the setup is shown in Fig. 2.

A typical real dolphin sonar signal was recorded, stored and used in the experiments as a simulated dolphin sonar signal. This simulated dolphin sonar signal was projected at the cylinders and a Gage 1012 Data Acquisition Board operating at a sampling frequency of 1 MHz was used to digitize

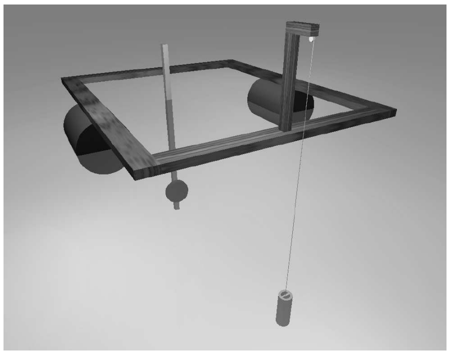

Figure 2: Illustration of the setup.

the echoes with a resolution of 16 bits. Each echo consisted of 1024 points with the time window placed so the echo from the target was present in the recorded signal. The emitted signal is shown in Fig. 3 and an example of recorded echoes from the standard and the +0.15 mm comparison target is shown in Fig. 4.

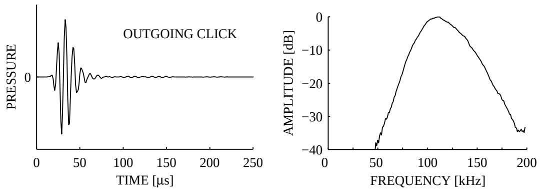

Figure 3: Waveform and spectrum of the projected simulated dolphin sonar signal.

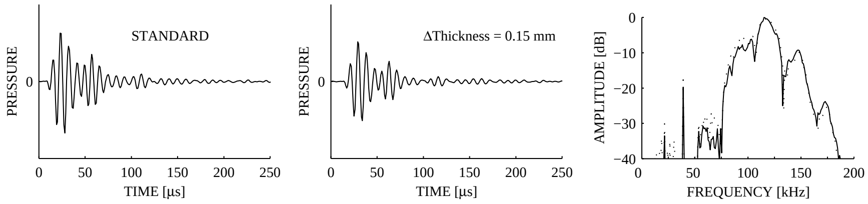

Figure 4: Echoes from the standard and the comparison cylinder that was 0.15 mm thicker than the standard cylinder. The waveforms are shown in the first two plots and the frequency spectra in the last plot. In the spectral plot the solid line is for the standard and the dotted line is for the comparison target.

The procedure for collecting the echoes was as following. the cylinder having the thinnest wall thickness was mounted at the end of the line and suspended in the water. Fifty echoes were then collected at a rate of two echoes per second. The cylinder was then removed from the water and unmounted from the line. This was performed for all ten cylinders. This session involving 500 echoes was repeated ten times giving a data base of 5000 echoes. This procedure was designed to reduce the risk of any potential features which are not related to the cylinders wall thicknesses (e.g. a fish swimming by, a cylinder hanging a little tilted etc.) being essential for classification of the echoes. Therefore, by using this procedure, hopefully only specific information from the cylinders will make it possible, when knowing the echoes in one session, to recognize echoes from the same cylinder from a different session.

PREPROCESSING

Features from the echoes were extracted using a combination of a matched filter, envelope detection, a gammatone filterbank, time integration and principal component analysis.

To find the beginning of each echo a matched filter, which was implemented as the time reversed version of the transmitted simulated dolphin click, was applied. The start of each echo was chosen as the peak of the envelope of the output from the matched filter (see e.g. [9]).

The gammatone filter is defined by its impulse response [1]

$$g(t) = at^{(n-1)} \exp(-2\pi bt) \cos(2\pi f_c t + \varphi), \; t > 0 \tag{1}$$

where b largely determines the duration of the impulse response and thus, the bandwidth of the filter; n is the order of the filter and it largely determines the slope of the skirts. When the order of the filter is in the range 3-5, the shape of the magnitude characteristic of the gammatone filter is very similar to that of the roex(p) filter commonly used to represent the magnitude characteristic of the human auditory filter [18]. The equivalent rectangular bandwidth (ERB) of the auditory filter is given by [11]

$$ERB = 24.75(4.37 f_{\mathrm{c}}/1000 + 1\mathrm{Hz}) \tag{2}$$

where f_{c} is center frequency of the filter. The center frequencies of the filters has been determined using Fay's modification of Greenwood's equation [10] for estimating cochlea frequency distributions along the basilar membrane

$$f_{\mathrm{c}}(x) = 0.008 f_{\max}(10^{2.1x} - 1.0) \tag{3}$$

where $f_{\max}$ is the maximal frequency perceived by the animal and x is the position of the filter on the basilar membrane expressed as the proportion between the distance from the basal end and the full length of the basilar membrane ($x = 0$ at the basal end and $x = 1$ at the apical end). A high-frequency boundary of 150 kHz for $f_{\max}$ was used to coincide with the bottlenose dolphin upper frequency limit of hearing [13].

If we choose $n = 4$ and f_{c}/b is large, which is the case here, then b and the 3-dB bandwidth, BW, of the filter are given by [17]

$$b = 1.019 \cdot ERB \tag{4}$$

$$BW = 0.887 \cdot ERB \tag{5}$$

The quality of a filter is defined as $Q = f_c/BW$ and by using the described function for center frequencies all filters have approximately a constant Q-value of 10 which is between the Q-values measured for the bottlenose dolphin using the two different techniques: Critical Bandwidth (CB) ($Q = 2.2$) [5] and Critical Ratio (CR) ($Q = 12.3$) [12].

The remaining constants a and φ in the gammatone filter are chosen as $a = 1$ as the amplification variable and $\varphi = 0$ as the phase variable.

The filterbank consists of N_f such gammatone filters and N_f is limited to 15 for computational reasons. The locations of the filters are chosen to be linearly spaced by the distance, $\mathrm{dx} = 0.0093$ with start and end locations corresponding to center frequencies of 80 kHz ($x = 0.86$) and 150 kHz ($x = 1$). The frequency response characteristic for the 15 gammatone filters are shown in Fig. 5.

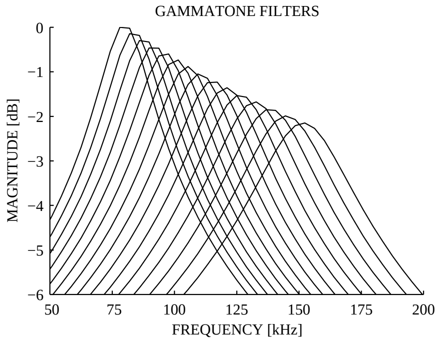

Figure 5: Frequency response for the gammatone filterbank.

The output of the filterbank when filtering the input signal with the bank of filters consists of N_f new signals, $y(n_\mathrm{f}, t)$, where n_f is the filter number.

Each of these signals was split up in $N_\mathrm{bin}(n_\mathrm{f})$ time bins of length $\delta_\mathrm{bin}(n_\mathrm{f}) = 1/f_c(n_\mathrm{f} \cdot \mathrm{dx})$ and the power was calculated in each time bin, $n_\mathrm{bin}(n_\mathrm{f})$ using

$$P(n_\mathrm{f}, n_\mathrm{bin}) = \frac{1}{\delta_\mathrm{bin}} \int_{(n_\mathrm{bin}-1)*\delta_\mathrm{bin}(n_\mathrm{f})}^{n_\mathrm{bin}*\delta_\mathrm{bin}(n_\mathrm{f})} y^2(n_\mathrm{f}, t)\mathrm{dt} \tag{6}$$

Only power values from time bins within the first 264 μs following the beginning of each echo were used for later analysis. A time window of 264 μs was chosen based on the dolphin's 264 μs integration time (see e.g. [6]). The power calculation in time bins are illustrated in Fig. 6. All power values were then combined into a single data vector P_all consisting of 439 values.

The complete data set consisting of all power value vectors were split in three sets: A training set, a validation set and a test set. The combined training-validation set consisting of the combination of the training set and the validation set was used to reduce the high dimensionality of the P_all vectors using principal component analysis (PCA) (see e.g. [20]). A singular value decomposition of the combined training-validation set was performed and only the ten highest eigenvalues were retained and used to project all

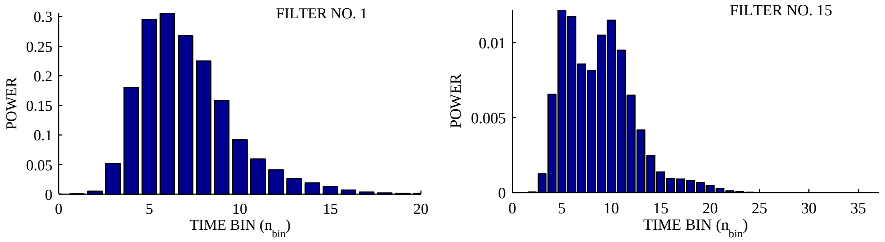

Figure 6: Illustration of power values calculated in time bins for the first filter and the last filter.

data down to a 10-dimensional feature vector, which means 10 features were extracted from each echo. The ten highest eigenvalues contained in average for the nine experiments 85 % of the variance. A limit of ten features was chosen for computational reasons to obtain a reasonable ratio between the number of training data and network complexity which is affected by the chosen number of features.

CLASSIFICATION

The extracted features were classified using a feed-forward net with a modified SoftMax [8] normalization as presented in [14] (see also, [2] [7]). The 2-layer feed-forward network with n_I inputs, n_H hidden neurons and $c-1$ outputs, where c is the number of classes, is defined by:

$$h_j(\mathbf{x}) = \tanh\left(\sum_{\ell=1}^{n_I} w_{j\ell}^I x_\ell + w_{j0}^I\right) \quad (7)$$

$$\phi_i(\mathbf{x}) = \sum_{j=1}^{n_H} w_{ij}^H h_j(\mathbf{x}) + w_{i0}^H \quad (8)$$

where $w_{j\ell}^I$, w_{ij}^H are the input-to-hidden and hidden-to-output weights, respectively. All weights are assembled in the weight vector $\mathbf{w} = \{w_{j\ell}^I, w_{ij}^H\}$. In order to interpret the network outputs as probabilities we used a *modified* normalized exponential transformation [14] similar to SoftMax [8],

$$\widehat{z}_i = \frac{\exp(\phi_i)}{\sum_{j=1}^{c-1} \exp(\phi_j) + 1}, \quad 1 \le i \le c-1, \quad (9)$$

$$\text{and} \quad \widehat{z}_c = 1 - \sum_{i=1}^{c-1} \widehat{z}_i, \quad (10)$$

where $\widehat{z}_i$'s are estimates of posterior probabilities.

The network was optimized using the maximum a posteriori technique, i.e., the cost function is the sum of the log-likelihood and a regularization term (prior).

$$C(\mathbf{w}) = S(\mathbf{w}) + R(\mathbf{w}, \kappa) \tag{11}$$

where $R(\mathbf{w}, \kappa)$ is a weight decay parameterized by a set of regularization parameters κ.

The full scheme for optimizing the network was presented at ICASSP 98 [14] but in this work only the weights and the regularization parameters were optimized using a second-order Gauss-Newton scheme based on the training set for the weights and a gradient descend scheme based on the validation set for the regularization parameters as described in [14]. For a more detailed description and use of outlier detection, see [14].

The number of inputs in the network correspond to the number of extracted features ($n_i = 10$), the number of hidden units, n_H was chosen to 5 and the number of classes c was chosen to 2 representing the standard cylinder and a comparison cylinder. Five hidden units was chosen to give an acceptable size of the network compared to the number of training data. Thus, the network initially consisted of $11 \cdot 5 + 6 \cdot 1 = 61$ free weights.

EXPERIMENTS

Ten experiments were performed testing the systems capability of discriminating between the standard cylinder and each of the ten comparison cylinders.

Data for the training set were chosen as all data from the first three sessions, data for the validation set came from the next three sessions while data from the last four sessions were used for the test set. The training, validation and test set for each experiment thus consisted of $50 \cdot 3 \cdot 2 = 300$, $50 \cdot 3 \cdot 2 = 300$ and $50 \cdot 4 \cdot 2 = 400$ data respectively since each set contained fifty echoes from each session from both the standard and a comparison cylinder.

Data from both the training set and the validation set were used for optimizing the neural network as described in [14] whereas data from the test set exclusively were used for testing the system.

RESULTS

Table 1 reports the performance of the system for different wall thickness differences between the standard and the comparison cylinders. The system is capable of discriminating between the standard and the comparison cylinder with a probability of correct classification, *pcc*, higher than 99% when the wall thickness difference is ≥ 0.15 mm. The system performance decreases as the wall thickness difference approaches zero and has a performance close to fifty percent (guessing) for a comparison cylinder expected to have the same wall thickness as the standard cylinder within the given mechanical tolerance. As can be seen in Table 1 retraining on the combined training and validation set using regularization parameters scaled by the increased number of training examples in the combined set improves the performance on the test set in most cases. Although increased performance is often observed using this technique it is not guaranteed. This can be seen in the experiment $\Delta WT = 0.00$ where a small decrease in performance is observed after retraining.

ΔWT [mm]	**Train.** [%]	**Val.** [%]	**Test** [%]	**Test after retrain** [%]
−0.31	100.0	100.0	100.0	100.0
−0.23	100.0	100.0	100.0	100.0
−0.15	100.0	100.0	99.5	99.5
−0.08	100.0	98.7	94.0	95.8
0.00	76.3	78.0	59.0	57.5
+0.08	100.0	93.7	83.5	96.0
+0.15	100.0	100.0	100.0	100.0
+0.23	100.0	100.0	99.8	99.8
+0.31	100.0	100.0	100.0	100.0

Table 1: Results from artificial experiment describing probability of correct classification for different wall thickness differences between the standard and the comparison cylinder.

Results from the artificial experiment are shown in Fig. 7 along with results from a similar experiment involving a false killer whale[1]. The performance of the system is seen to be comparable with the performance of the dolphin with a slightly higher probability of correct classification for the artificial system. The confidence intervals are found using the standard formula for calculating confidence intervals for a binomial distribution.

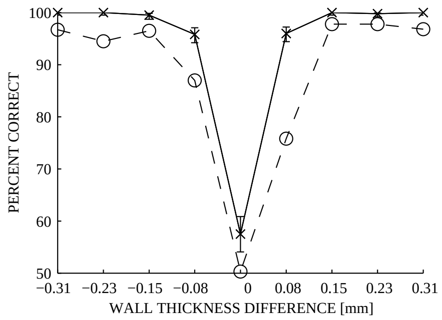

Figure 7: Results from artificial experiment (x-markers connected with full line including 90%-confidence intervals) compared with results from a similar experiment involving a false killer whale (o-markers connected with dashed line).

[1]This experiment has not yet been published.

DISCUSSION AND CONCLUSION

The results from the experiment shows that the described system can discriminate very accurately between the standard cylinder and the comparison cylinders differing more than or equal to 0.15 mm in wall thickness. This small wall thickness difference has been detected under conditions where the targets have been approximately five meters away and subject to wind and wave motions.

Comparing the results from the artificial experiment with results from a similar experiment involving a false killer whale shows that the artificial system can produce performance results which are comparable to results obtained by the whale. It should be emphasized however that although there are similarities in the two experiments differences are also present. In the whale experiment the whale is trained to remember the standard cylinder and during the experiment choose the standard when presented for the standard and a comparison cylinder. This task thus demands the dolphin to remember the standard cylinder over a period of time. The dolphin might also be distracted or simply unconcentrated during the experiment. Such psychological factors are important to have in mind when comparing the results and are possibly some of the reasons why the whale has a slightly lower performance than the artificial system.

Having this in mind the results of this study using a combination of the matched filter technique, the gammatone filterbank, time integration, principal component analysis and neural networks for discriminating cylinders with different wall thicknesses has demonstrated its potential.

ACKNOWLEDGMENTS

The authors would like to express their appreciation to Dr. Robert Gisiner, Office of Naval Research, for allowing us to use data from the experiment involving the *pseudorca crassidens*, to Dr. Herbert Roitblat, University of Hawaii, for many inspiring discussions and to Dr. Paul Nachtigall, University of Hawaii, for making the collaboration possible. Research was supported by the Danish Research Councils through the Danish Computational Neural Network Center (CONNECT) and the THOR Center for Neuroinformatics.

REFERENCES

[1] A. Aertsen and P. Johannesma, "Spectro-temporal receptive fields of auditory neurons in the grassfrog. 1. characterisation of tonal and natural stimuli," **Biol. Cybern.**, vol. 38, pp. 115–135, 1980.

[2] L. Andersen, J. Larsen, L. Hansen and M. Hintz-Madsen, "Adaptive regularization of neural classifiers," in J. Principe, ed., **IEEE Workshop on Neural Networks for Signal Processing VII**, IEEE, Piscataway, NJ, USA, 1997, pp. 24–33.

[3] W. Au, **The Sonar of Dolphins**, New Jersey, USA: Springer Verlag, 1993.
[4] W. Au, L. Andersen, A. Rasmussen, H. Roitblat and P. Nachtigall, "Neural network modeling of a dolphin's sonar discrimination capabilities," **J. Acoust. Soc. Am.**, vol. 98, pp. 43–50, 1995.
[5] W. Au and P. Moore, "Critical ratio and critical bandwidth for the atlantic bottlenose dolphin," **J. Acoust. Soc. Am.**, vol. 88, pp. 1635–1638, 1990.
[6] W. Au, P. Moore and D. Pawloski, "Detection of complex echoes in noise by an echolocating dolphin," **J. Acoust. Soc. Am.**, vol. 83, pp. 662–668, 1988.
[7] C. Bishop, **Neural Networks for Pattern Recognition**, Oxford, UK: Oxford University Press, 1995.
[8] J. Bridle, "Probabilistic interpretation of feedforward classification network outputs with relationships to statistical pattern recognition," in **Neurocomp. - Algor. , Arch. and Appl.**, Berlin, Germany: Springer-Verlag, vol. 6, pp. 227–236, 1990.
[9] W. Burdic, **Underwater Acoustic System Analysis**, New Jersey, USA: Prentice Hall, 1991.
[10] R. Fay, "Structure and function in sound discrimination among vertebrates," in D. Webster, R. Fay and A. Popper, eds., **The Evolutionary Biology of Hearing**, New York, USA: Springer-Verlag, pp. 229–263, 1992.
[11] B. Glasberg and B. Moore, "Derivation of auditory filter shapes from notched-noise data," **Hearing Research**, vol. 47, pp. 103–138, 1990.
[12] C. Johnson, "Masked tonal threshold in the bottlenose porpoise," **J. Acoust. Soc. Am.**, vol. 44, pp. 965–967, 1968.
[13] C. Johnson, "Auditory masking on one pure tone by another in the bottlenose porpoise," **J. Acoust. Soc. Am.**, vol. 49, pp. 1317–1318, 1971.
[14] J. Larsen, L. Andersen, M. Hintz-Madsen and L. Hansen, "Design of robust neural network classifiers," in **ICASSP'98**, Seattle, USA, 1998, vol. 2, pp. 1205–1208.
[15] P. Moore, H. Roitblat, R. Penner and P. Nachtigall, "Recognizing successive dolphin echoes with an integrator gateway network," **Neural Networks**, vol. 4, pp. 701–709, 1991.
[16] P. Nachtigal, "Odontocete echolocation perfomance on object size, shape and material," in R. Busnel and J. Fish, eds., **Animal Sonar**, New York, USA: Plenum, pp. 71–95, 1980.
[17] R. Patterson, J. Holdsworth, I.Nimmo-Smith and P. Rice, "Svos final report: The auditory filterbank," Tech. Rep. 2341, APU, 1988.
[18] R. Patterson and B. Moore, "Auditory filters and excitation patterns as representations of frequency resolution," in B. Moore, ed., **Frequency Selectivity in Hearing**, London, UK: Academic Press Limited, pp. 123–177, 1986.
[19] R. Patterson, K. Robinson, J. Holdsworth, D. McKeown, C. Zhang and M. Allerhand, "Complex sounds and auditory images," in **Advances in the Biosciences**, UK: Pergamon Press, vol. 83, pp. 426–446, 1992.
[20] B. Ripley, **Pattern Recognition and Neural Networks**, Cambridge, UK: Cambridge University Press, 1996.
[21] H. Roitblat, P. Moore, P. Nachtigal, R. Penner and W. Au, "Natural echolocation with an artificial neural network," **Neural Networks**, vol. 1, pp. 239–248, 1989.