Proceedings of the Institute of Acoustics

AUTOMATIC DOLPHIN WHISTLE DETECTION, EXTRACTION, ENCODING, AND CLASSIFICATION

C. Sturtivant & S. Datta

Underwater Acoustics Group, Electronic & Electrical Engineering, Loughborough University

1. INTRODUCTION

The Underwater Acoustics Group at Loughborough University has developed software for quantitative comparison of frequency modulated tonal sounds, specifically for those produced by dolphins [1,2,3]. The software can identify parts of recordings containing whistle-like sounds, extract their frequency-time-intensity contours, and then apply automatic pattern recognition techniques to classify them against previous whistles. These techniques have the benefits of being both objective and quantitative, and provide a probability that any candidate whistle belongs to each existing class, or to some new class.

This identification technique relies on dolphins producing distinctive identifying whistles. This 'signature' whistle theory was proposed over thirty years ago [4], and suggested that the majority of dolphin whistles carried the identity of the vocalising animal. Although recently questioned [5], the more current proposal still maintains that whistles carry identity information, but in the wider context of groups of animals. Such identifying whistles have been found for several dolphin species as well as for the killer whale [6,7,8,9,10], and so potentially have applications to a wide number of toothed cetaceans.

Identification of groups of wild dolphins can be most beneficial during studies of behavioural changes. The traditional use of photographic identification for in the field [11,12] has two main drawbacks. Firstly, it is reliant on good visibility, which limits its use to above the surface considering the clarity of British coastal water, and also is only of use during the day. Secondly, since dolphins spend very little time at the surface to breathe, an observer must be close enough to take clear pictures. When one is attempting to record changes in behaviour, any observer close enough to the dolphins for photo ID work might easily affect the very behaviour that is being observed. Thus, one of the main advantages with acoustic identification is that it does not require an observer in the vicinity. The presence of a hydrophone is far smaller and is less intrusive than a human observer on a boat with a camera.

2. EQUIPMENT AND DATA

Recordings were made on the Dutch research vessel Tridens as part of a parallel project [13,16,17] named CETASEL (Commission of European Communities contract number AIR3-CT94-2423). Between 15:06 and 15:17 on 10th October 1996 whistles were noted from three groups of common dolphins. These recordings were analysed to attempt to distinguish between the three groups. The signals from a trawl-mounted hydrophone were preamplified (with a Benthos AQ4/AD743) before being sent via a coaxial cable some 450 metres back to a ship, where acoustic and visual observations were logged. An R-DAT recorder (Sony TCD-D7) with a 22 kHz bandwidth was used to record the resulting signals.

Two intervals of relative quiet split the recorded whistles into three, labelled as periods A, B, and C. These periods contained 7, 12, and 30 extractable whistles respectively, although the signal to noise of some whistles was too low for analysis. All the whistles were classified according to the automated classification procedure for each period as explained below.

3. WHISTLE DETECTION

Dolphin whistles have two features that distinguish them from background noise on a spectrogram: their narrow frequency bandwidths, and a relatively stable central frequency over short times. A discrete fast Fourier transform (FFT) was carried out on the data to convert the signal into the time/frequency domain. For viewing purposes, a transform partition size of 256 samples (corresponding to 5.8 ms) produced good results, but for the detection routine a smaller partition size of 32 samples (corresponding to 0.73 ms) was used. This briefer time resulted in frequency bins with a width of 689 Hz, so that whistles (which change only slowly in frequency) changed infrequently between adjacent bins.

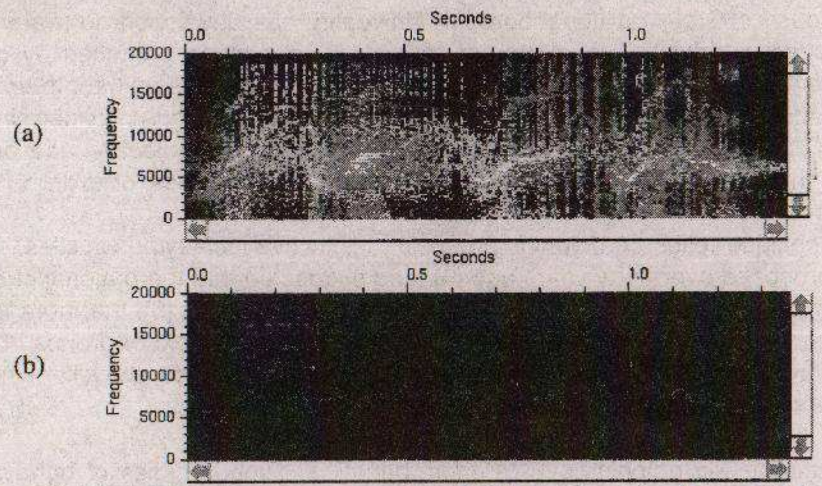

Figure 1 - Spectrogram of a dolphin whistle and clicks: (a) original data; (b) data after filtering

In order to reduce the contribution to the signal made by impulsive noises (such as echolocation clicks), a filtering technique was used to enhance signals with narrow frequency bandwidths. A typical whistle contour masked by noise from echolocation clicks is shown (Figure 1a). Echolocation clicks have a much broader frequency range than whistles, so a click removal method was devised whereby the average energy between time partitions was normalised. This calculation is described in Equations 1 and 2.

$$\overline{E}(t) = \frac{1}{N}\sum_{k=0}^{N-1}\left|f(k,t)\right|^2 \qquad (1)$$

$\overline{E}(t)$ represents the average energy in time partition t, and $\left|f(k,t)\right|^2$ the energy in frequency bin k for time partition t. The normalised energy value E_n was set to a suitable constant value, and the new filtered transform $f'(k,t)$ calculated from Equation 2.

$$f'(k,t) = \sqrt{\frac{E_n}{\overline{E}(t)}}\, f(k,t) \qquad (2)$$

However, some remnant of the echolocation click was often present in the filtered transform, since the click spectrum was rarely uniform. This prompted a modified algorithm, where the average energy was calculated only from the values in the m surrounding frequency bins (Equation 3).

$$\overline{E}(k,t) = \frac{1}{2m+1}\sum_{l=k-m}^{k+m}\left|f(l,t)\right|^2 \qquad (3)$$

This new average, dependent on the central frequency bin, was used to calculate the filtered transform (Equation 4)

$$f'(k,t) = \sqrt{\frac{E_n}{\overline{E}(k,t)}} f(k,t) \quad \text{(4)}$$

This filtering produced a noise 'trough' around the whistle, since the high intensities in the whistle would have a large effect on the average energy calculation for surrounding frequencies (Figure 1b).

In order to remove any background or remaining transient components to the signal, two averages were taken of the resulting filtered signal using an exponential decay (Equation 5). The background noise was assessed with a large value for α giving a "half-life" of several seconds, and the averaged current signal by using smaller values for α giving a half-life of a few milliseconds.

$$x'_{ave} = \alpha x_{ave} + (1-\alpha)x, \text{ where } 0 \le \alpha \le 1 \quad \text{(5)}$$

The instantaneous (filtered) signal spectrum could then be calculated as the difference between the averaged current signal and the background noise. Frequency regions with discrete energy peaks were then flagged as possible whistles. If a peak was maintained for more than a threshold minimum time in the same of adjacent frequency bins, it was flagged as a potential whistle.

Since the software was not fast enough to carry out whistle detection and store flagged data to disk, a log was made of those times when a detection was made. The times durations indicated would begin before the initial detection was made and end after the whistle was flagged as finished, in order that none of the whistle contour was cut off. A second pass of the program over the recording was necessary to sample the data onto the computer's hard disk at the times specified by the log.

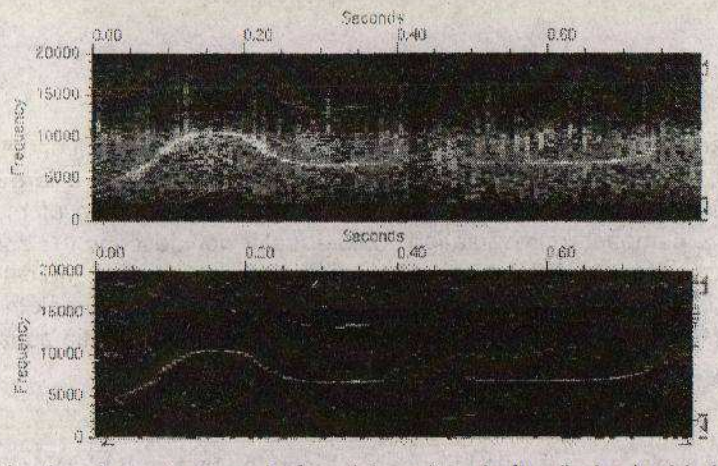

Figure 2: Example signal spectrogram before (above) and after (below) whistle enhancement.

4. WHISTLE EXTRACTION

The method for extracting the whistles' time-frequency-intensity contours uses some of the filtering techniques laid out above. The fast Fourier transform is taken of signals in which whistles have been detected, and the edge detection filter (Equations 3 & 4) is applied to the resulting spectrogram to reduce broadband noises such as echolocation clicks (Figure 2). A contour following routine is used that attempts to follow a smooth path through the spectrogram, whereby a higher intensity is required to suddenly alter the direction of the traced contour.

The contour is subdivided into a number of segments in order to reduce the data requirements for comparisons. Each segment indicates periods when the contour is generally rising, falling, or flat in frequency, or a temporary silence. The time-frequency information within each of these segments is approximated by a quadratic equation using a least squares fitting routine [2]. Thus the characteristic 'shape' of the contour is kept with a marked reduction in data, although currently intensity information is not preserved (Figure 3).

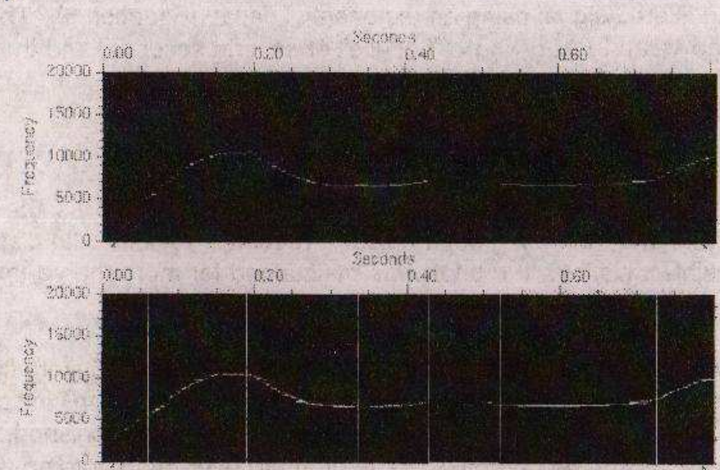

Figure 3: An extracted whistle contour showing segments and fitted curves

5. WHISTLE CONTOUR ENCODING

Attempting to compare one whistle with another by using the simple sequence of time-frequency pairs extracted by the previous algorithm clearly would involve comparison of a large amount of data, since whistle durations are typically between 0.2 and 3 seconds. Observation of the 'shape' of the whistle suggested that a more compact representation could equally well describe the salient characteristics. An algorithm was adopted whereby the whistle was split up into segments, indicating whether the whistle was 'rising', 'flat', or 'falling' in frequency with time, or 'blank' indicating a break in the contours. The data points contained in each of these segments was represented as a quadratic equation of the familiar form shown in Equation 6.

$$y(x) = a_0 + a_1 x + a_2 x^2 \quad \text{......(6)}$$

The origin for x and y is placed at the first frequency bin in the first time partition of the segment, allowing the curves of similar segments to be compared. In general, the quadratic equation would not exactly match all points on the extracted whistle contour, so a least-squares fitting routine was used. The equation was solved by using singular value decomposition (SVD) [14] on the points within each segment.

6. CLASSIFICATION PROCEDURE

Whistle classification was based around two features: overall contour shape, and detailed contour structure differences. Hidden Markov models, or HMMs, were employed to represent segment sequences for each contour class. Details of HMM theory and suggested algorithms can be found in

Rabiner's tutorial paper [15]. The HMM can be used to determine the probability of producing a specific segment sequence given a class, when it has been trained to represent that class.

A separate comparison was made in addition to the HMM method, based on three measures between corresponding segments for pairs of contours. The quadratic equation parameters allow calculation of average differences in frequency, frequency slope, and rate of change of slope for pairs of segments, and standard deviations of these values could be calculated over all contours within a class. Since these measures form a Gaussian distribution within a class, the degree to which a candidate whistle is representative of a class can be found as a percentage based on comparisons with those whistles already contained in that class.

Differences in average frequency can be calculated by integrating over the two segments that are to be compared (Equation 7).

$$\delta_0 = \int_{t=0}^{T} \left(f_1(t) - f_2(t)\right) dt \qquad (7)$$

Rather than calculating this integral, the result can be found from the quadratic parameters from Equation 6. Similarly, δ_1 and δ_2, indicating the average difference in rate of change of frequency (frequency slope) and rate of change of frequency slope, can also be calculated from the quadratic parameters.

The product of this similarity percentage and the probability from the hidden Markov model was used to calculate the probability that a whistle from the specified class would match the candidate whistle. The probability for membership of each class for a specific whistle could then be calculated using Bayes's theorem for conditional probabilities, using an empirical probability that the whistle might belong to an as yet undiscovered class. Any candidate whistle could then be placed either in the class with the highest membership probability, or in a new class if so indicated.

7. WHISTLE GROUP CLASSES

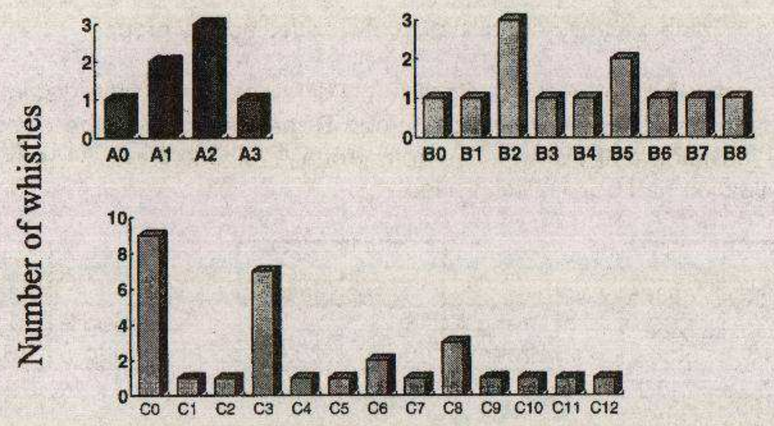

Figure 4 — Automatic classification of whistles from the three groups.

The whistles from each period in the recordings were termed whistle 'groups', such that group A contained all whistles from period A, etc. The classification procedure described above was applied to each group of whistles. Thus, within each group, the whistles were assigned to one of a number of different classes. These classes and membership are shown graphically in Figure 4.

The classes within each group are labelled with the group letter followed by the class number. Several classes within each group contained just one whistle, and it is possible that some were due to 'aberrant' whistle from dolphins, i.e. a whistle type other than that individual's 'signature' whistle. Alternatively this species' whistles may be variable and carry no information unique to individuals or groups. Comparing whistles between groups can test this last hypothesis.

8. CROSS-GROUP COMPARISON

Classification was attempted on contours from one group with classes characterised for another (Table 1). The probabilities for class membership were calculated for each whistle-class pair, and then summed by group and class to give the expected class membership for each group based on the classification parameters. Chi-squared analysis was made on class-group pairs with the null hypothesis of no difference in whistle classes between groups (Table 2). The resulting significance probability could then be used to determine whether two groups of dolphins produced the same types of whistle.

	A0	A1	A2	A3	B0	B1	B2	B3	B4	B5	B6	B7	B8
Group A	1	2	3	1	0	0	0	0	(2)	0	0	0	0
Group B	0	0	0	0	1	1	3	1	1	2	1	1	1

	A0	A1	A2	A3	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Group A	1	2	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Group C	0	2	1	0	9	1	1	7	1	1	2	1	3	1	1	1	1

	B0	B1	B2	B3	B4	B5	B6	B7	B8	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Group B	1	1	3	1	1	2	1	1	1	3	0	1	0	0	0	0	0	2	0	0	0	0
Group C	0	0	9	1	0	3	1	0	0	9	1	1	7	1	1	2	1	3	1	1	1	1

Table 1: Cross-classification of whistles across groups.

Chi-squared analysis revealed group A had a significantly different whistle type distribution from either of the other two groups. None of the whistles from group B matched any of the classes from group A, although the software indicated that two whistles from group A fell into class B4 (bracketed in Table 1), which on further investigation had been misclassified.

	A Classes	B Classes
Group A	5.72	1.90
Group B	2.46	9.56

$\chi^2 = 3.784$, 1 d.f., $p = 1.75\%$

	A Classes	C Classes
Group A	5.72	2.35
Group C	6.25	25.09

$\chi^2 = 7.874$, 1 d.f., $p = 1.39\%$

	B Classes	C Classes
Group B	9.56	7.02
Group C	15.39	25.09

$\chi^2 = 1.844$, 1 d.f., $p = 17.45\%$

Table 2: Expected number of whistles per class per group based on membership probabilities.

Although three whistles out of 30 from group C were matched by classes formed from group A whistles, none of those from group A fell into group C classes. If this fact is taken in combination with the chi-squared probability of 1.39%, this table indicates group A whistles are significantly different to those in group C.

Similar analysis of classes and whistles from groups B and C indicated some difference between groups, but not a signification one. Further analysis was conducted on just those classes that contained more than one whistle, termed the 'major' classes for a group, and thus eliminating classes consisting of aberrant whistles (Table 3).

	B2	B3	B5	B6	C0	C2	C3	C6	C8
Group B	3	1	2	1	3	1	0	0	2
Group C	9	1	3	1	9	1	7	2	3

Table 3: Whistle comparison for 'major' classes in groups B and C.

Classes B2 and C0, B3 and C2, and B5 and C8 all contained the same whistles, suggesting that these class pairs were identical, although classes C3 and C6 contained no whistles from the other group. One possible explanation is that the two groups of dolphins were recorded simultaneously for a time, and whistles from C3 and C6 belonged solely to the second of the two groups.

9. CONCLUSIONS

An analysis software tool has been developed that to a large extent can automate the task of pattern recognition of dolphin whistles. Several species of dolphins use identifying whistles, giving this tool a wide applicability to studies of reoccurrence of individual groups. When applied to a species that had not previously been known for its identifying whistles, the software was able to encode and classify the whistles from common dolphins by contour shape. Using these classes, it was shown that one of the three groups contained whistles significantly different from the other two, and that the other two groups had been recorded simultaneously for part of the time. When whistles from the earlier of the groups were removed from the data, there were significantly different whistles from the latter group. Thus, the whistles from all three groups could be used to determine group identities and to separate them in time.

This result shows that a combination of hidden Markov modelling and statistical modelling of dolphin whistle time-frequency contours can be used to characterise different whistle types. A data reduction technique that preserved the frequency-time 'shape' of the whistle was employed, and since this did not obstruct the group identification, appears to be appropriate to the situation. Although limited in the number of whistles they contain, these results strongly suggest that the common dolphin employs whistles which can be used to identify individual groups.

10. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the help of the CETASEL project members for providing the data for this research, and the aid of Kristin Kaschner, David Goodson, and Professor Bryan Woodward. Funding for this project was provided by the U.K. Department of the Environment under contract number CR 0129, and the Ministry of Agriculture, Fisheries, and Foods under contract CSA 2270. Much of the preliminary work on this project would not have been possible without the help of several trainers and oceanariums, especially Mr Peter Bloom of Flamingo Land, U.K., and Mats Amundin and Susanne Hultman of Kolmårdens Djurpark, Sweden.

11. REFERENCES

[1] C R STURTIVANT & S DATTA, 'Techniques to isolate dolphin whistles and other tonal sounds from background noise', *Acoustics Letters* **18(10)**, 189–193 (1995).

[2] C R STURTIVANT & S DATTA, 'The isolation from background noise and characterisation of bottlenose dolphin (*Tursiops truncatus*) whistles', *Journal of the Acoustical Society of India* **23(4)**, 199–205 (1995).

[3] C R STURTIVANT & S DATTA, 'An acoustic aid for population estimates', *European Research on Cetaceans — 11* (1997).

[4] M C CALDWELL & D K CALDWELL, 'Individualized whistle contours in bottlenosed dolphins (*Tursiops truncatus*)', *Nature* (London) **207**, 434–435 (1965).

[5] B MCCOWAN & D REISS, 'Quantitative comparison of whistle repertoires from captive adult bottlenose dolphins (Delphinidae, *Tursiops truncatus*): a re-evaluation of the signature whistle hypothesis', *Ethology* **100**, 194–209 (1995).

[6] M C CALDWELL & D K CALDWELL, 'Statistical evidence for individual signature whistles in Pacific whitesided dolphins, *Lagenorhynchus obliquidens*', *Cetology* **3**, 9 pp. (1971).

[7] M C CALDWELL, D K CALDWELL, & J F MILLER, 'Statistical evidence for individual signature whistles in the spotted dolphin, *Stenella plagiodon*', *Cetology* **16** (1973).

[8] J K FORD & H D FISHER, 'Killer whale (*Orcinus orca*) dialects as an indicator of stocks in British Columbia', *International Whaling Commission Scientific Committee Report* SC/JN81/KW8 (1981).

[9] M E DAHLHEIM, 'Signature information in killer whale calls', *Whalewatcher* **15(1)**, 12–13 (1981).

[10] J K FORD & H D FISHER, 'Group specific dialects of killer whales (*Orcinus orca*) in British Columbia', In *Communication and behavior of whales* (R. Payne, ed.), AAAS selected Symposia Series, Westview, Boulder, Colorado, pp. 129–161, (1983).

[11] R S WELLS, A B IRVINE, & M D SCOTT, 'The social ecology of inshore odontocetes', In *Cetacean Behavior: Mechanisms and Functions*, L.M. Herman (ed.), Robert E. Krieger, Florida, pp. 263–317, (1980).

[12] B WÜRSIG & M WÜRSIG, 'The photographic determination of group size, composition, and stability of coastal porpoises (*Tursiops truncatus*)', *Science* **198**, 755–756, (1977).

[13] P R CONNELLY, A D GOODSON, K KASCHNER, P A LEPPER, C R STURTIVANT, & B WOODWARD, 'Acoustic techniques to study cetacean behaviour around pelagic trawls', *Proceedings of ICES Conference*, Baltimore, USA, 25 September – 1 October 1997.

[14] W H PRESS, B P FLANNERY, S A TEUKOLSKY, & W T VETTERLING, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press.

[15] L R RABINER, 'A tutorial on hidden Markov models and selected applications in speech recognition', *Proceedings of the IEEE* **77(2)**, 257–285 (1989).

[16] A D GOODSON, D NEWBOROUGH, & B WOODWARD, 'Set gillnet acoustic deterrents for harbour porpoises, *Phocoena phocoena*: improving the technology', *Proceedings of ICES Conference*, Baltimore, USA, 25 September – 1 October 1997.

[17] A D GOODSON, M AMUNDIN, R H MAYO, D NEWBOROUGH, P A LEPPER, C LOCKYER, F LARSEN, & C BLOMQVIST, 'Aversive sounds and sound pressure levels for the harbour porpoise (*Phocoena phocoena*): an initial field study', *Proceedings of ICES Conference*, Baltimore, USA, 25 September – 1 October 1997.

especially Mr Peter Bloom of Flamingo Land, U.K., and Mats Amundin and Susanne Hultman of Kolmårdens Djurpark, Sweden.

11. REFERENCES

[1] C R STURTIVANT & S DATTA, 'Techniques to isolate dolphin whistles and other tonal sounds from background noise', *Acoustics Letters* **18(10),** 189–193 (1995).

[2] C R STURTIVANT & S DATTA, 'The isolation from background noise and characterisation of bottlenose dolphin (*Tursiops truncatus*) whistles', *Journal of the Acoustical Society of India* **23(4),** 199–205 (1995).

[3] C R STURTIVANT & S DATTA, 'An acoustic aid for population estimates', *European Research on Cetaceans — 11* (1997).

[4] M C CALDWELL & D K CALDWELL, 'Individualized whistle contours in bottlenosed dolphins (*Tursiops truncatus*)', *Nature* (London) **207,** 434–435 (1965).

[5] B MCCOWAN & D REISS, 'Quantitative comparison of whistle repertoires from captive adult bottlenose dolphins (Delphinidae, *Tursiops truncatus*): a re-evaluation of the signature whistle hypothesis', *Ethology* **100,** 194–209 (1995).

[6] M C CALDWELL & D K CALDWELL, 'Statistical evidence for individual signature whistles in Pacific whitesided dolphins, *Lagenorhynchus obliquidens*', *Cetology* **3,** 9 pp. (1971).

[7] M C CALDWELL, D K CALDWELL, & J F MILLER, 'Statistical evidence for individual signature whistles in the spotted dolphin, *Stenella plagiodon*', *Cetology* **16** (1973).

[8] J K FORD & H D FISHER, 'Killer whale (*Orcinus orca*) dialects as an indicator of stocks in British Columbia', *International Whaling Commission Scientific Committee Report* SC/JN81/KW8 (1981).

[9] M E DAHLHEIM, 'Signature information in killer whale calls', *Whalewatcher* **15(1),** 12–13 (1981).

[10] J K FORD & H D FISHER, 'Group specific dialects of killer whales (*Orcinus orca*) in British Columbia', In *Communication and behavior of whales* (R. Payne, ed.), AAAS selected Symposia Series, Westview, Boulder, Colorado, pp. 129–161, (1983).

[11] R S WELLS, A B IRVINE, & M D SCOTT, 'The social ecology of inshore odontocetes', In *Cetacean Behavior: Mechanisms and Functions*, L.M. Herman (ed.), Robert E. Krieger, Florida, pp. 263–317, (1980).

[12] B WÜRSIG & M WÜRSIG, 'The photographic determination of group size, composition, and stability of coastal porpoises (*Tursiops truncatus*)', *Science* **198,** 755–756, (1977).

[13] P R CONNELLY, A D GOODSON, K KASCHNER, P A LEPPER, C R STURTIVANT, & B WOODWARD, 'Acoustic techniques to study cetacean behaviour around pelagic trawls', *Proceedings of ICES Conference*, Baltimore, USA, 25 September – 1 October 1997.

[14] W H PRESS, B P FLANNERY, S A TEUKOLSKY, & W T VETTERLING, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press.

[15] L R RABINER, 'A tutorial on hidden Markov models and selected applications in speech recognition', *Proceedings of the IEEE* **77(2),** 257–285 (1989).

[16] A D GOODSON, D NEWBOROUGH, & B WOODWARD, 'Set gillnet acoustic deterrents for harbour porpoises, *Phocoena phocoena*: improving the technology', *Proceedings of ICES Conference*, Baltimore, USA, 25 September – 1 October 1997.

[17] A D GOODSON, M AMUNDIN, R H MAYO, D NEWBOROUGH, P A LEPPER, C LOCKYER, F LARSEN, & C BLOMQVIST, 'Aversive sounds and sound pressure levels for the harbour porpoise (*Phocoena phocoena*): an initial field study', *Proceedings of ICES Conference*, Baltimore, USA, 25 September – 1 October 1997.