Echolocation sound sources in the Harbour Porpoise - a wavelet approach to classification

A. D. Goodson, O. Farooq, S. Datta

Underwater Acoustics Group, Electronic and Electrical Engineering Dept., Loughborough University, LE11 3TU, UK. a.d.goodson@lboro.ac.uk, o.farooq@lboro.ac.uk, s.datta@lboro.ac.uk

Abstract

While testing acoustic signals designed to mitigate harbour porpoise by-catch mortality in commercial gillnets, a large volume of echolocation activity from a pair of free-swimming harbour porpoises (Phocoena phocoena) was recorded for behavioural analysis. The need to identify the individual vocalising animal in these recorded sequences requires an effective classification tool. This species transmit a stereotypical narrow-band poly-cyclic echolocation signal and as some progressive phase and amplitude modulation effects are also observed during trains of pulses these signals offer few obvious features to aid discrimination. The available data set included a sub-set of 'on-axis' examples, where the vocalising individual had been identified visually. Conventional analytical methods offered limited success and the objective of an automatic classifier remains to be achieved. A new classification method, based on a novel high resolution wavelet transform technique, appears to be more robust, and this approach encourages further development. The presence of an initial very stable component within the first few cycles of the porpoise signals is confirmed, and the modulation effects observed within these pulses appear to be a controlled phase coding artifact created by the signal generation mechanism.

1. Introduction

Echolocation evolved as the primary prey detection and interception mechanism for dolphins and porpoises. Dolphins in general employ a high power wide-band short echolocation pulse with reported source levels (SL) up to 220 dB_{rms} re 1 µPa at 1m [1], although some smaller dolphin species, (e.g. Commerson's dolphin *Cephalorhynchus Commersonii*) have opted for low power narrow band signals similar to those employed by the harbour porpoise.

Figure 1. The harbour porpoise (Phocoena phocoena) - the echolocation signal source.

The harbour porpoise (*Phocoena phocoena*, Figure 1) transmits a narrow band sonar pulse centred around 130-150 kHz with SLs in the order of 160 dB re 1 µPa at 1 m [2-6] but, unlike a dolphin, this animal does not emit frequency modulated 'communication' whistles. The high frequency pulsed sound is emitted from the forepart of the head, from a fatty tissue structure known as the 'melon'. This organ functions as a graded-index waveguide as the melon tissues comprise both short and long chain lipids fractions, which are distributed to provide a nonlinear sound velocity structure [7, 8]. Progressively higher values of the sound speed surround a low velocity core material, with some marked asymmetry in the vertical plane that may beneficially minimise near-field effects [9, 10]. The physiology that appears to define the bandwidth of the porpoise pulse has been modelled [11, 12] and the lip-like small dorsal bursæ (MLDB) sited just above the nasal plug are structures clearly implicated as the impulsive source of sound production [13]. As two similar sized bursæ pairs are present displaced symmetrically from the median, a question exists as to whether these two structures operate independently, or co-operate, during echolocation signal generation.

2. Signal Sources and Data Capture

Studies, of a male and a female harbour porpoise reacting to sounds introduced into their enclosure, were carried out at the Fjørd and Bælt Centre at Kerteminde in Denmark. It was noted that, when the sound signals ceased at the end of each test, the animals would proceed to interrogate the position of the source using echolocation. These studies also demonstrated that synthesised signals closely emulating the porpoise's own echolocation pulses would stimulate 'investigative' echolocation behaviour.

High resolution examples of mid-water 'on-axis' sonar signals were captured using a Tektronix TDA420 digital oscilloscope (200 MHz, 100 Msamples s^{-1}). Echolocation pulse trains were also recorded using a high speed instrumentation tape recorder (Racal Store 4 DS - 30"/s, 150 kHz bandwidth). These signals were recorded from a small 12.5 mm ball hydrophone (Sonar Products Ltd - HS150) positioned immediately adjacent to the stimulation source transducer.

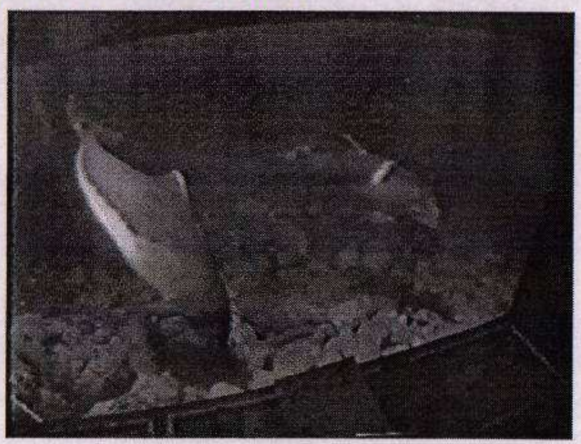

Figure 2. Short range echolocation activity. Porpoises foraging for fish between the rocks.

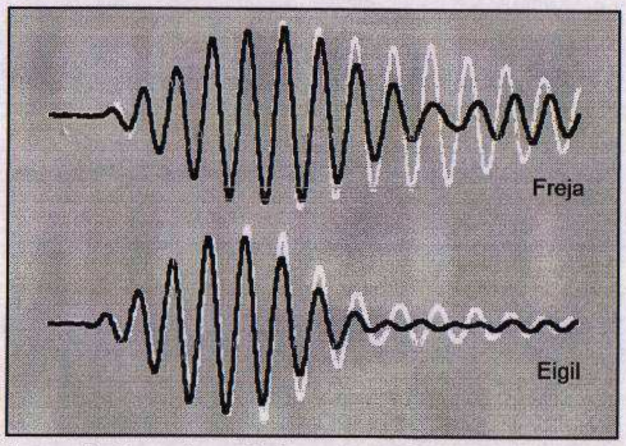

Figure 3. Porpoise echolocation pulses recorded from the male 'Eigil' and female 'Freja' showing the similar centre frequency with some amplitude modulation. The first 4 cycles do not include modulation.

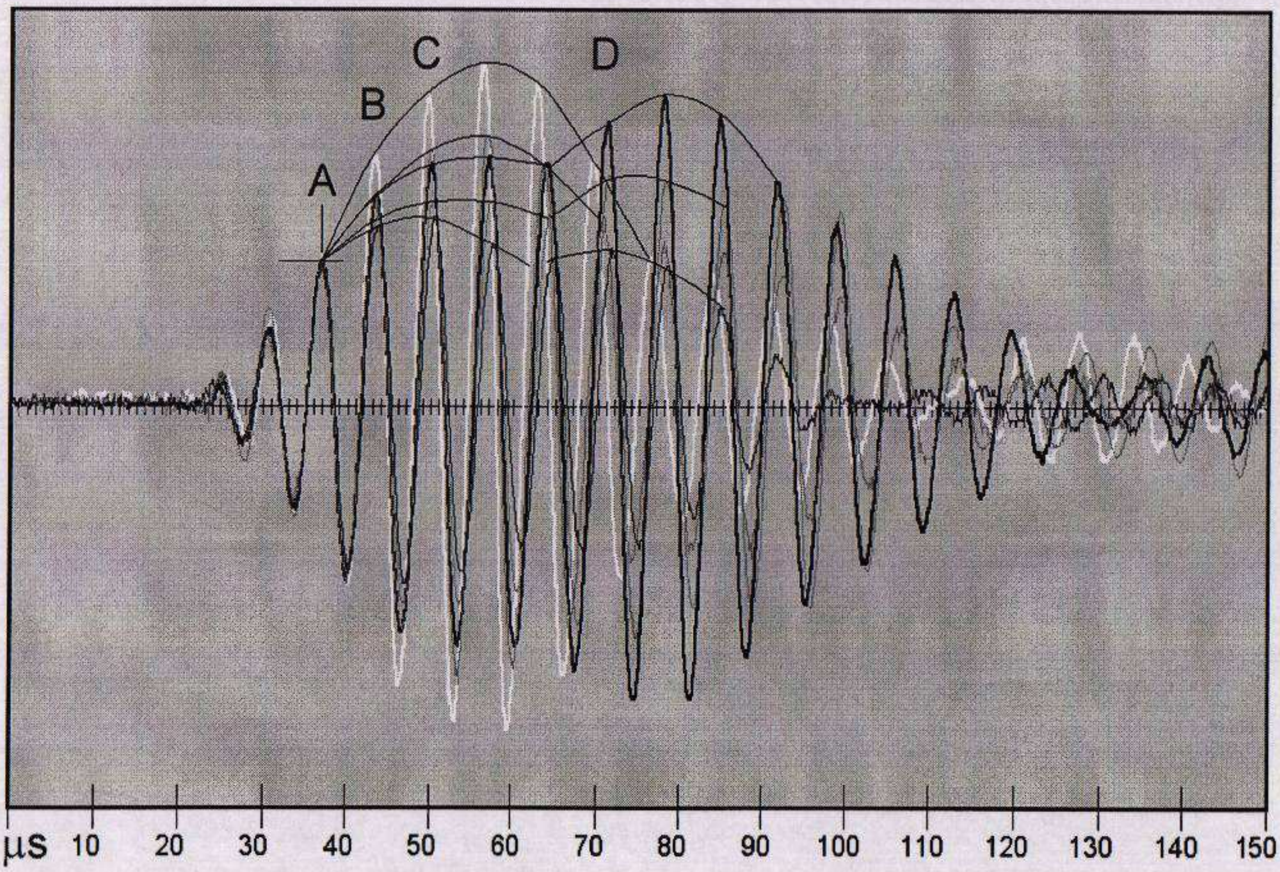

*Figure 4. Eigil - 5 click samples from the male porpoise overlaid after normalising their respective amplitudes at the 3rd cycle (**A**). The signals increase in amplitude to the 4th cycle (**B**) after which the envelope either increases or decreases (**C**) suggesting summation with a similar signal component delayed some 3.5 to 4 cycles (**D**).*

The porpoise echolocation signals were also processed in real time using an envelope 'click detector' (Loughborough University). This made these ultrasonic pulses both audible and recordable on the sound track of the video recorders used to log the position and orientation of the animals within the enclosure. The recording equipment was synchronised during these trials using standard EBU (VITC and LTC) time code.

The two harbour porpoises studied, a male and female, (Figure 2), were rescued in April 1997 from a herring 'pound net' fish trap at Kørsor, West Sealand. The animals were estimated to be 2-3 years old at that time. By June 1999, around the time of the recordings they had grown significantly (female: 47.5 kg, length=148 cm; male: 40.8 length=138 cm). These porpoises were held for study in Denmark under a special licence issued to aid research into methods of reducing porpoise mortalities in bottom-set fishing nets. The EC supported project 'EPIC' [14] involving partners based in Denmark, Sweden and the UK included tests of acoustic signals for 'aversive' or 'attractive' characteristics.

A preliminary analysis of the data recordings revealed that significant amplitude and phase modulation is normally present in these signals. The modulation effects starting to appear within each pulse after a short but fairly constant delay which could not be explained by external multi-path effects (Figures 3 and 4).

3. Analysis Methods

Before attempting to analyse the volume of echolocation sequences recorded on the instrumentation recorder a means of identifying the echolocation signals originating from each animal was required.

The set of over-sampled digital recordings captured via the oscilloscope was first examined in the laboratory for features that would allow these data to be partitioned between the two animals. A sub-set of these data files included the identity of the source animal as reported by an observer at the time of recording. The classification of these signals proved more difficult than expected. This is because both animals produced very similar signal spectra, and the variations in pulse shape and phase which occurred during trains of pulses helped to mask the more obvious cues. Source levels were not considered a useful discriminator, although in some cases the transmission loss could be estimated as path lengths could be determined from the video record. The intended application required that the individual porpoises should be identifiable from acoustic characteristics, so that echolocation behaviour could be correctly assigned during analysis. Peak frequencies in the power spectra were examined as body size was expected to affect this. Juvenile porpoises [5] have been shown to operate at higher frequencies than adults [4], and the larger animal (in this case the female) was expected to transmit at a frequency slightly below that of the smaller male. The application of a simple Fast Fourier Transform to these signals did not, however, discriminate reliably. The variable phase shift within the individual pulses may partially explain this. The novel use of a Wavelet Transform approach gave much better results, and this is the technique summarised here.

4. The Wavelet Transform Classifier

A Wavelet Transform (WT) [15-17] performs a constant 'Q' (Quality Factor) analysis of a signal by projecting it on a set of basis functions whose scale varies with frequency (Equation 1). A mother wavelet is shifted and scaled to generate additional wavelet functions that are orthogonal to each other. This WT signal processing approach decomposes the signal into low pass and high pass components, and then down-samples it by 2. If required, the inverse transform can perform a reconstruction. The filters used for decomposition permit perfect reconstruction, and the WT functions appear orthogonal with finite impulse responses. These are called conjugate mirror filters.

A wavelet is a function $\psi \in L^2$ (3) (i.e. a finite energy function) with zero mean and is normalised ($\|\psi\| = 1$). A family of wavelets can be obtained by scaling ψ by s and translating it by u.

$$\psi_{u,s}(t) = s^{-1/2}\psi\left(\frac{t-u}{s}\right) \tag{1}$$

The Continuous Wavelet Transform (CWT) of a finite energy signal $\mathbf{f(t)}$ is given by:

$$CWTf(u,s) = \int_{-\infty}^{+\infty} f(t).s^{-1/2}.\psi^{*}\left(\frac{t-u}{s}\right)dt \tag{2}$$

where $\psi^{*}(.)$ is the complex conjugate of $\psi(.)$. The above equation can be viewed as convolution of the signal with dilated band-pass filters.

The Discrete Wavelet Transform (DWT) of a signal $f[n]$ with period N is computed as:

$$DWTf[n,a^j] = \sum_{m=0}^{N-1} f[m].a^{-j/2}.\psi^*\left(\frac{m-n}{a^j}\right) \tag{3}$$

where m and n are integers. The value of a is equal to 2 for a dyadic transform.

Using a digital signal processing approach the DWT decomposes the signal into low pass and high pass components and then down-samples these by 2; the inverse transform performs the reconstruction. Figure 4 shows

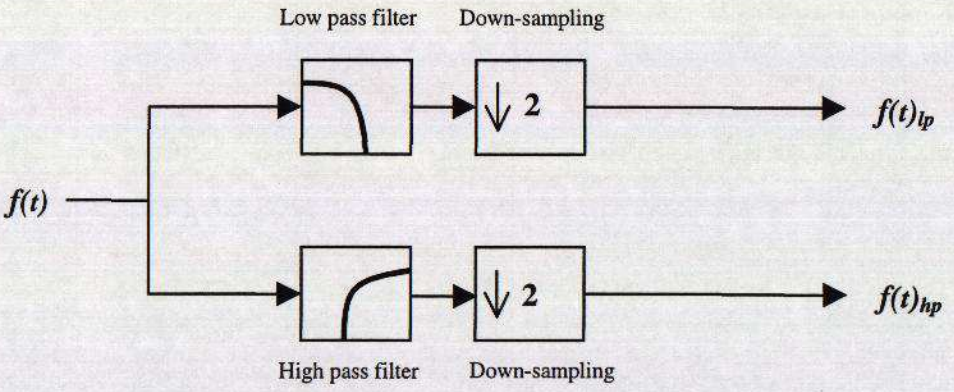

Figure 4. First level decomposition of signal $f(t)$ by DWT

the decomposition of the signal $f(t)$ into two components, $f(t)_{lp}$ a low pass version and $f(t)_{hp}$ a high pass. A second level of decomposition is applied to the $f(t)_{lp}$ signal only, and the $f(t)_{hp}$ is left untouched. Thus a DWT gives a left recursive binary tree structure, where the left branch represents the lower frequency band. A more general form is the Wavelet Packet Transform (WPT) [16]. This tries to decompose the lower, as well as the higher, frequency bands, thereby giving a balanced binary tree structure. For the purpose of classification, the required features might come from some specific frequency bands, which sometimes may not be extracted by DWT (because it only splits the lower frequency band). The drawback of WPT is that it contains a high level of redundant information. Thus for the purpose of feature extraction, we propose the use of the Admissible Wavelet Packet Transform (AWPT). By using the AWP transform, any selected frequency band can be split to enhance discrimination for classification purposes.

As the sampling frequency of the Porpoise clicks was 10 MHz, the Nyquist signal bandwidth in the data is 0-5 MHz. By applying the decomposition by DWT, this band is split into two: a low pass band with frequency 0-2.5 MHz, and a high pass band of 2.5-5 MHz. The second level of decomposition splits the lower band again into two (i.e. 0-1.25 MHz and 1.25-2.5 MHz). By applying six successive levels of decomposition, the seven bands obtained are: 0-78 kHz, 78-156 kHz, 156-312 kHz, 312-625 kHz, 625-1,250 kHz, 1.25-2.5 MHz and 2.5-5 MHz. The last three frequency bands carry no significant energy from the biological signal, and hence could be discarded. The second band has maximum energy, and hence it may contain more discriminatory information. Keeping this in mind, this band is further decomposed and split into two bands (78-117 kHz and 117-156 kHz). Energy of the wavelet coefficients in these four bands is calculated, and these are used as features for the male and female classification process. A Linear Discriminant Analysis (LDA) classifier [18] was implemented and the results obtained using these features are shown in Table 1.

LDA is a tool used for multi-group data classification and dimensionality reduction. It tries to minimise the ratio of within-class scatter to between-class scatter and thereby attempts to achieve maximum separability. A within-class scatter matrix defines the scatter of samples around their respective class mean. Between-class scatter matrix defines the spread of the mean vectors around the global mean. LDA tries to separate the different group data by forming a linear decision boundary between them.

5. Results

	Training		Testing	
	Male	Female	Male	Female
Male (Case 1)	9	0	5	1
Female (Case 1)	0	6	0	3
Male (Case 2)	8	1	6	0
Female (Case 2)	0	6	0	3
Male (Case 3)	9	0	5	1
Female (Case 3)	0	6	0	3

Table 1. Confusion Matrix of the classification results

As the data set included 11 digital recordings where the animal's identity had been indicated by observers, these examples were divided into 'training' and 'test' data sets. To increase the size of these small data sets, additional samples were created by the addition of white Gaussian noise to produce degraded signal-to-noise ratio versions (SN=10 dB and 15 dB). This gave a total of 24 male and 9 female clicks. In order to generalise the results, three different mutually exclusive combinations of training and test examples were examined.

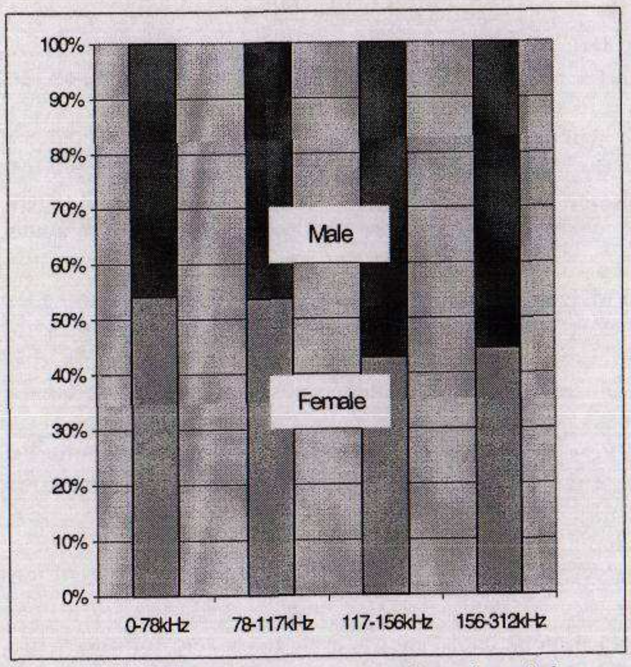

Figure 6. AWPT frequency bins. Case 1 training set comparing male/female energy-averaged for illustration purposes.

Figure 6 illustrates the approach, using averaged energy data from the case 1 training set in each frequency bin. The LDA classifier correctly assigned the female test examples in all of these combinations. However, this classifier consistently rejected a single degraded SN example (15 dB signal-noise ratio) derived from a male, and incorrectly assigned this to be a female in both the training and test combinations.

6. Discussion and Conclusions

The Admissible Wavelet Packet Transform provides a classification technique that may be implemented in real time and automated. In this application the technique appears to be acceptably robust and may be expected to work with relatively noisy signals containing significant phase and amplitude modulation. The approach may also offer advantages when attempting to classify target echo data.

The basic objective, to achieve a satisfactory discrimination between these two animals using only frequency domain features that could be extracted from individual echolocation pulse waveforms, was broadly achieved through this Wavelet processing approach. However, the test methods we employed have examined 'whole pulse' waveforms. It seems clear that if the initial 4 to 5 cycles of the harbour porpoise waveform which precede any significant phase or amplitude modulation effect can be isolated, then this signal segment may permit alternative methods to be reconsidered. The porpoise echolocation waveforms we have examined confirm the earlier observations of Kamminga and Wiersma [4]. They showed that Harbour porpoise echolocation signals frequently include a phase modulated component, and we concur with their speculation that this effect could be caused by the addition of a delayed signal. However, the observed delay is remarkably stable and too short to be caused by an external multipath reflection via the water surface. Within the melon waveguide structure internal reflections do not appear possible [9, 10], and the mechanical adjustment of a 'reflector', suitably positioned within the head of the animal, that can switch on, or off, a delay path of some 26 μs (4 cycles @ 150 kHz) implies controlled movement in the order of 15 to 20 mm (30-38 mm/2 estimated from the sound velocity in similar tissue). As the morphology does not support such an adjustable reflection capability we can reject both internal and external reflections as the probable cause. We prefer the simpler alternative that both MLDB structures can operate in combination to initiate closely spaced pulse-pairs with a minimum separation of at least 20 μs. The signal sequences we have observed from these two animals suggests that, in addition to this delay, some fine control of phase may be possible. If this is so then such phase-coded sonar transmissions may offer some target classification advantages.

7. Acknowledgements

The support of the European Commission DG XIV is gratefully acknowledged. Omar Farooq acknowledges funding support from the Commonwealth Commission (British Council). Special thanks are due to our project EPIC partners and to all of our colleagues who participated in the original field trials. We especially thank the porpoise trainers, Sabrina Labberte, Kirstin Anderson, and the staff of the Fjørd and Baelt Centre at Kerteminde. We appreciated the help and co-operation offered by the Zoology Department, and by post-graduate students, of the University of Southern Denmark (Odense). We also wish to record our thanks to the two enthusiastic harbour porpoises Eigil and Freja who provided the acoustic data.

References

[1] Au WWL. Echolocation signals of the Atlantic bottlenose dolphin (*Tursiops truncatus*) in open waters, in *Animal Sonar Systems,* Plenum Press, New York, 1980, pp. 251-282.

[2] Dubrovskii NA, Krasnov PS and Titov AA. On the emission of echolocating signals by the Azov Sea harbour porpoise. *Soviet Physics-Acoustics* 1971; **16**(4): 444-448.

[3] Mohl B and Andersen S. Echolocation: high frequency component in the click of the harbour porpoise (*Phocoena phocoena* L.). *Journal of the Acoustic Society of America* 1973; **54**:1368-1373.

[4] Kamminga C and Weirsma H. Investigations on Cetacean Sonar II, acoustic similarities and differences in odontocete sonar signals. *Aquatic Mammals* 1981; **8**(2): 73-94.

[5] Goodson AD and Sturtivant C. Sonar characteristics of the harbour porpoise (*Phocoena phocoena*) source levels and spectrum. *ICES Journal of Marine Science* 1996; **53**: 465-472.

[6] Goodson AD, Kastelein RA and Sturtivant CR. Source Levels and echolocation signal characteristics of juvenile Harbour porpoises (*Phocoena phocoena*) in a pool, in *Harbour porpoises - Laboratory studies to reduce bycatch.* De Spil, Woerden, NL, 1995, pp. 41-53.

[7] Norris KS and Harvey GW. Sound transmission in the porpoise head. *Journal of the Acoustic Society of America* 1974; **56**: 659-664.

[8] Cranford TW, Amundin M and Norris KS. Functional morphology and homology in the odontocete nasal complex: Implications for sound generation. *Journal of Morphology* 1996; **228**: 223-285.

[9] Goodson AD, Flint J, Pomeroy S and Cranford TW. Bio-sonar characteristics of the harbour porpoise (*Phocoena phocoena*). in: Underwater Acoustics ECUA2000. European Commission, Brussels. ISBN 920828-9530-0. pp. 479-484.

[10] Goodson AD, Flint J and Cranford TW. The harbour porpoise (*Phocoena phocoena*) – Modelling the sonar transmit mechanism. in: Advances in the study of echolocation in bats and dolphins. Chicago University Press, in press.

[11] Goodson AD and Datta S, Investigating the Sonar Signals of the Harbour Porpoise, *Phocoena Phocoena*, *Proceedings of the Journal of the Acoustical Society of India.* 1995; **23**(4): 205-211.

[12] Goodson AD. A narrow band bio-sonar: Investigating echolocation in the harbour porpoise, *Phocoena phocoena. Proceedings of the Institute of Acoustics*, 1997; **19**(9): 19-28.

[13] Cranford TW, Van Bonn WG, Ridgway SH, Chaplin MS and Carr JA. Functional morphology of the dolphin's biosonar signal generator studied by high-speed video endoscopy. *Journal of Morphology* 1997; **232**(3): p243.

[14] Lockyer C, Amundin M, Desportes G, Goodson AD and Larsen F. EPIC - Elimination of Harbour porpoise Incidental Catch. Final Report to the European Commission DG XIV on Special Study Project DG XIV 97/0006: 2000; 246pp., submitted.

[15] Rioul O and Vetterli M. Wavelet and signal processing. *IEEE Signal Processing Magazine* October 1991: 14-38.

[16] Mallat S. A wavelet tour of signal processing. Academic Press, San Diego, California, USA, 1998.

[17] Farooq O and Datta S. Dynamic feature extraction by wavelet analysis. *Proceedings of the 6th International Conference on Spoken Language Processing*, Beijing, China, 2000; 4: 696-699.

[18] Fukunaga K. Introduction to statistical pattern recognition, Academic Press, San Diego, California, USA, 1990.