THE PERFORMANCE OF THE SPUD ALGORITHM DETECTING HARBOUR PORPOISE (*PHOCOENA PHOCOENA*) ECHOLOCATION CLICKS

EJ Harland Chickerell Bioacoustics, Weymouth, Dorset, UK.

1 INTRODUCTION

Automated self-contained harbour porpoise (*Phocoena phocoena*) echolocation click detectors, such as the POrpoise Detector (POD) unit available from Chelonia [1], have been used for some years to detect when animals are present in a particular area. They can also give an indication of the behaviour of the animals from the patterns of echolocation pulses. These units have found widespread use in a number of applications.

However, there has also been some doubt expressed about the meaning of the data collected and it would appear that at times there can be a significant variation in achieved performance between units deployed in the same area. This variation appears to be due to a number of factors, including manufacturing spread, deployment problems and operator settings.

The original POD design was produced nine years ago. Although there have been a number of later versions it is still based on comparing the outputs of a number of analogue filters to determine the presence of an echoloation click. The original proto-POD used a four filter system. It is believed the current POD uses a two filter system.

The aim of the the Simple Porpoise Underwater Detector (SPUD) work was to attempt to make a detection system that would have the following characteristics

- No operator controls
- Good manufacturing reproducibility
- All processing internal to unit
- Easy deployment
- Recovery aids to assist relocation
- Small size

Spud is a sub-set of a more advanced concept, the Complex Research Underwater Detector (CRUD). CRUD aims to detect and classify all echolocation and similar clicks while SPUD is highly optimised to detect only the echolocation clicks from the harbour porpoise.

This paper describes the origins of the SPUD algorithm, how it operates and the results of testing with acoustic data containing harbour porpoise and other clicks

2 PROTO-POD

The prototype of POD, known as proto-POD was designed by the author in 1996 and is used here as an example of detecting echolocation clicks using analogue filter methods. The block diagram of proto-POD is shown in figure 1. The spectrogram of the harbour porpoise echolocation clicks are shown in the top right sub-pane and the waveform in the bottom sub-pane. They are typically 150 micro-seconds long and with a bandwidth of 30 kHz centred on 130 KHz.

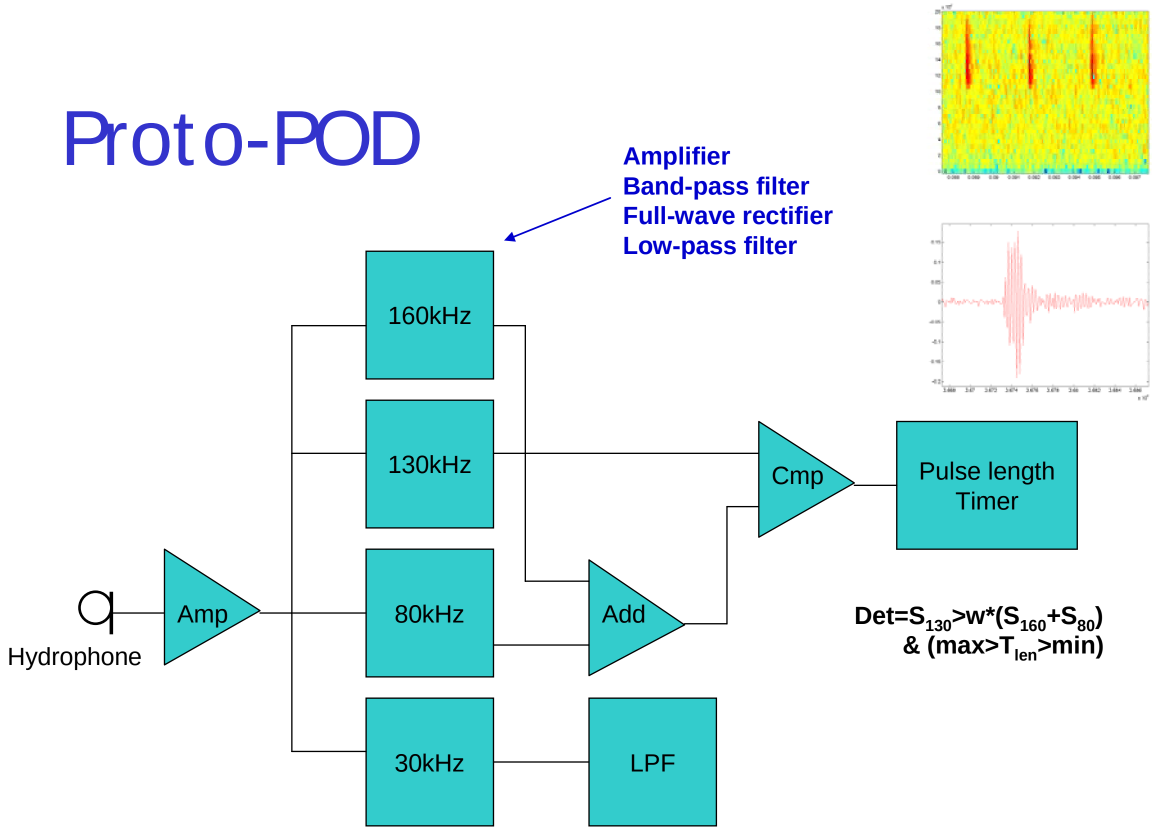


Figure 1. Proto-POD block diagram

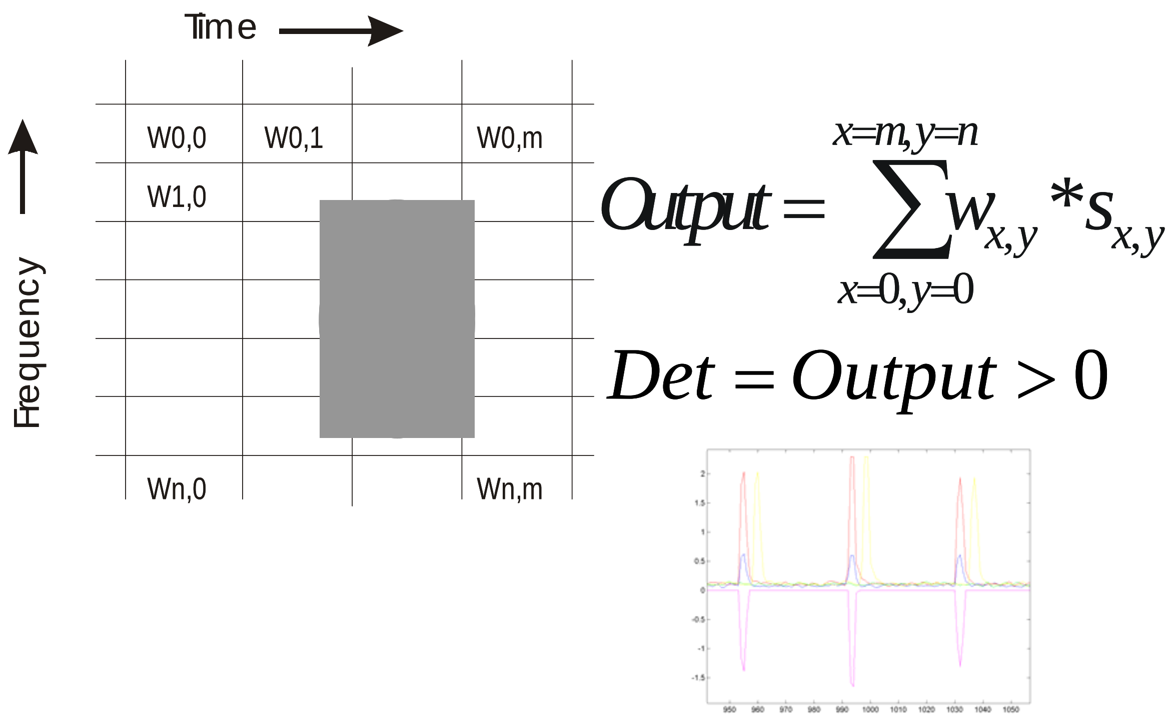
The signal from the hydrophone is amplified and then applied to four bandpass filters. The energy at 130 kHz is compared with the energy in bands above and below the bandwidth occupied by the echolocation pulse, in this case 80 and 160 kHz. A fourth bandpass filter at a low frequency, in this example 30 kHz, is used to assess ambient noise levels and can be used to de-sensitise the detection comparison to reduce false alarms under high ambient noise conditions such as a passing boat. The processing block consists of a fourth-order Butterworth bandpass filter, a full-wave rectifier and a low-pass filter. A set of analogue comparators and a gate array are then used to implement the detection test.

For best performance it was necessary to match the time repsonse of the three filters so that the signals being compared were time aligned. However, Butterworth analogue filters also have sidelobes in both the frequency and time domain. For best detection performance it was necessary to match these sidelobes across all three filters as well as the main filter response, and this proved very difficult. At high signal levels when the animals were very close it was possible for each click to be detected several times in thre time sidelobes. This problem was made worse if the signal through any of the filters was sufficiently strong to cause clipping. Ideally a fast-acting, peak responding, automatic gain control should be used. As with all analogue systems dynamic range is a problem. Proto-Pod achieved around 60dB range between just detecting a signal and the inception of clipping.

3 SPUD

SPUD is based on the concept of spectrogram correlation proposed by Mellinger and Clark [2, 3]. The incoming acoustic data is sampled at 400 kHz and then passed through a 64 point FFT to give a 32 point spectrogram. Data is collected for 1 second and then processed to detect the presence of echolocation clicks. The current version of SPUD is implemented using MATLAB (Ver 5.3) so it is not possible to process in real-time. The MATLAB code can run with either a National Instruments fast A/D collecting batched data and then processing off-line, or it can be used to process data from WAV files. All of the testing described in this paper was carried out by reading pre-recorded data from WAV files. The data came from a variety of sources and used various sample rates. All data was converted to 400 kHz sample frequency using the CoolEdit program.

Figure 2 shows the processing method used by SPUD.



$$Output = \sum_{x=0,y=0}^{x=m,y=n} w_{x,y} * s_{x,y}$$

$$Det = Output > 0$$

Figure 2. SPUD processing diagram

Cells in the spectrogram around the echolocation pulse are allocated weighting factors which are positive for cells where the pulse is expected to be and negative for those cells which should not contain energy from the pulse. A detection occurs when the sum of the weighted energy levels in the cells is positive. The sub-pane shows SPUD in operation. The top two traces are the outputs from two sub-sums of the weighted cells. The higher amplitude trace is the sum of the positive weighted cells while the lower amplitude trace is the sum of the negative weighted cells. The yellow trace is the time delayed sum. The lower trace is the inverted detection function.

One of the challenges of spectrogram correlation methods is choosing the correct set of weighting coefficients. A variety of examples of echolocation clicks are needed to optimise these coefficients and it may even be necessary to have multiple reference sets to suit different acoustic environments or variations in the calls of the animals. In the case of the harbour porpoise the echolocation call is sufficiently different from other echolocation calls to be heard in North-East European waters that designing the reference set is comparatively straightforward and one set should suffice all conditions and call variants. The SPUD coefficient set was chosen by manually inspecting two datasets. One was from two captive animals at the Fjord and Baelt Centre in Denmark, the other from animals in the wild in the Baltic Sea.

The output from the spectrogram correlator processing was then scanned to produce a pulse list containing time of pulse, amplitude of pulse and width of pulse. The MATLAB routine then outputs this list to a CSV file readable by EXCEL. The pulse sequence is then scanned to build pulse trains. The statistics of the pulse trains can be used to further classify the calls and also to estimate the number of echolocating animals present. Au *et al* and Teilman *et al* [4, 5] both noted that the harbour porpoise prefer an inter-click interval around 60 ms, although intervals as low as 30ms and as high as 200ms have been observed.

Initially this pulse train processing was implemented manually in EXCEL but as confidence grew a MATLAB routine was written to perform the processing automatically. Although this part of the work is not yet complete, it is sufficiently complete to allow testing of the complete SPUD concept.

The pulse train processing uses a tree search algorithm to find connected pulses. Starting at pulse P0, the algorithm searches forward until the next pulse is found that has amplitude and width similar to P0 and within time limits set by the characteristics of the animal. This pulse is designated P1. The algorithm then carries out a search constrained in time, amplitude and width for P2. This search then continues until no more pulses are found. As the search progresses a running confidence factor is calculated based on how well the next pulse fits the prediction from previous pulses. If this confidence factor drops below a threshold the train is deemed to have ended. The algorithm can cope with up to two consecutive missed pulses. If there are multiple pulses within the constrained search window the pulse giving the highest confidence factor is chosen.

As the pulse train is built the algorithm needs to calculate the statistics of variations in the inter-click interval in order to build confidence in the classification decision. At present, development is still continuing on this part of the SPUD system. A much larger dataset than is currently available is needed to complete this work.

The final output from the completed SPUD algorithm will be the presence of animals. The user will have available a list of times when animals were detected, an estimate of the number of animals and the confidence of the classification decision. Once more is known about the way the animals use their echolocation capability during various activities it may also be possible to provide the user with animal activity information. Also provided to the user will be an estimation of detection range as acoustic conditions vary so that the coverage likely to have been achieved can be documented.

4 TESTING SPUD

Testing was carried out in two phases. The first phase tested the spectrogram correlation in some detail while the second phase tested the pulse train processing. Two MATLAB programs were used. The first read WAV files of raw acoustic data in blocks of one second of data, carried out the spectrogram correlation, displayed the results in a series of MATLAB figures and then scanned the processed data to build the pulse list. The second MATLAB programme read the pulse list and then attempted to build the pulse trains.

Datasets used to test the data came from captive and wild harbour porpoise. In addition, two more datasets were used to test the false positive rate. One datset came from captive bottlenose dolphins (*Tursiops truncatus*) while the other dataset came from an area with a high level of non-cetacean biological activity and a very high probability that no animals were present. It is believed that the clicks in this datset came from crustaceans.

For each dataset the individual pulses were documented manually and assigned a probable source. The datsets were then processed by the SPUD MATLAB programs and the SPUD decision for each pulse compared with the manual decision. This is a time consuming process so not all of the available data has been processed to date.

The pulse train testing used just one dataset which contained an unknown number of animals in the wild. Because of a number of difficulties in the testing of this algorithm the work has not progressed as fast as was initially hoped. However, the current algorithm is able to assemble pulse trains and it was possible to assess its performance.

5 RESULTS

The spectrogram processing was tested using four datasets. Two of the datasets were individual harbour-porpoise, one dataset was captive bottlenose dolphins and the fourth dataset was crustacean clicks.

The results from the two harbour porpoise datasets were:

Number of harbour porpoise clicks determined visually:	587
Number detected by SPUD	585
Number of other pulse detected visually	5
Number of other pulses miss-classified by SPUD:	0

The two missed pulses were both too weak to exceed the chosen threshold. The five 'other pulses' were random noise pulses of unknown origin and SPUD correctly determined they were not harbour porpoise clicks.

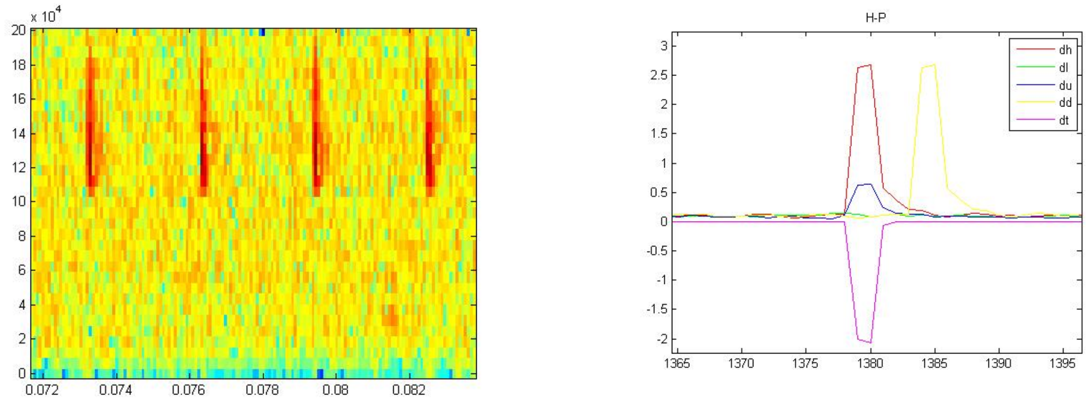


Figure 3. Processing a harbour porpoise click

Figure 3 shows the spectrogram used by SPUD for the spectrogram processing stage. The right-hand pane shows the outputs from the various stages of the processing. Red is the sum of the block of cells corresponding with the echolocation clicks, green is the sum of cells lower in frequency, blue is the sum of cells higher in frequency, yellow is the time offset sum and purple is the inverted detection decision.

The results from the bottlenose dolphin dataset were:

Number of bottlenose dolphins clicks determined visually:	176
Number of bottlenose clicks miss-classified by SPUD:	17
Number of false positives	137

The high number of false positives were caused by a noise source within the tank that may have been due to a pump running. It generated a wideband repetitive pulse and a broad band of highly variable noise in the region of 120-160 kHz. Unfortunately this noise band corresponds to the frequency range of the harbour porpoise echolocation click. Because of the variability of this signal it is inevitable that at times it will meet all the criteria set by SPUD in classifying the pulses.

The dolphin clicks mis-classified by SPUD were investigated in some detail. The two primary causes were the spectral characteristics of the reverberation tail on the main pulse and spectral distortion of multipath arrivals. In both cases a dip in the spectral characteristics had been introduced by the acoustic environment and so SPUD classified them as harbour porpoise clicks. This dataset had originally been recorded with a sample frequency of 320 kHz and had been re-sampled at 400 kHz for compatibility reasons. It meant that there was no acoustic noise above 160 kHz so the comparison with the upper region was not possible. In most of these false positive cases the presence of energy in the upper band would have prevented the detection.

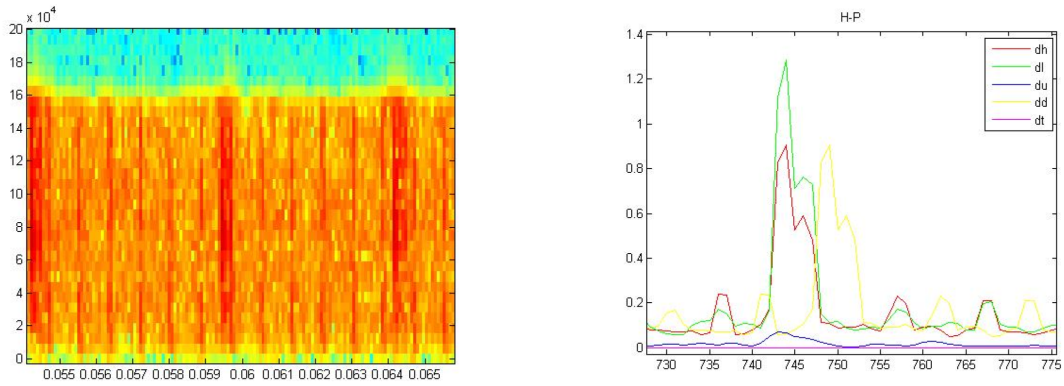


Figure 4. Processing bottlenose dolphin clicks

Figure 4 shows the spectrogram of the bottlenose clicks. The faster, weaker regular clicks are believed to be mechanical noise. The right-hand pane again shows the outputs from various parts of the processing, but this time it can be seen that the purple lower trace does not go negative, indicating that SPUD rejected the click as not being from a harbour porpoise.

The results for the Crustacean click dataset were:

Number of clicks determined visually:	211
Number of clicks mis-classified by SPUD:	1

Close examination of the one mis-classification showed that acoustic propagation had introduced a broad dip in the spectrum in the 60-80 kHz region, which meant that SPUD determined it to be from a harbour porpoise.

Figure 5 shows the spectrogram for a typical crustacean click. These are very loud clicks and this example was recorded close to a harbour wall so there are six possible paths detectable. The right-hand pane shows SPUD processing this signal. Close inspection of the spectrogram shows that the third arrival has a peak around 120 kHz and lower energy at 80 kHz. This is reflected in the levels in the right-hand pane and in this case the pulse came close to satisfying the detection criteria.

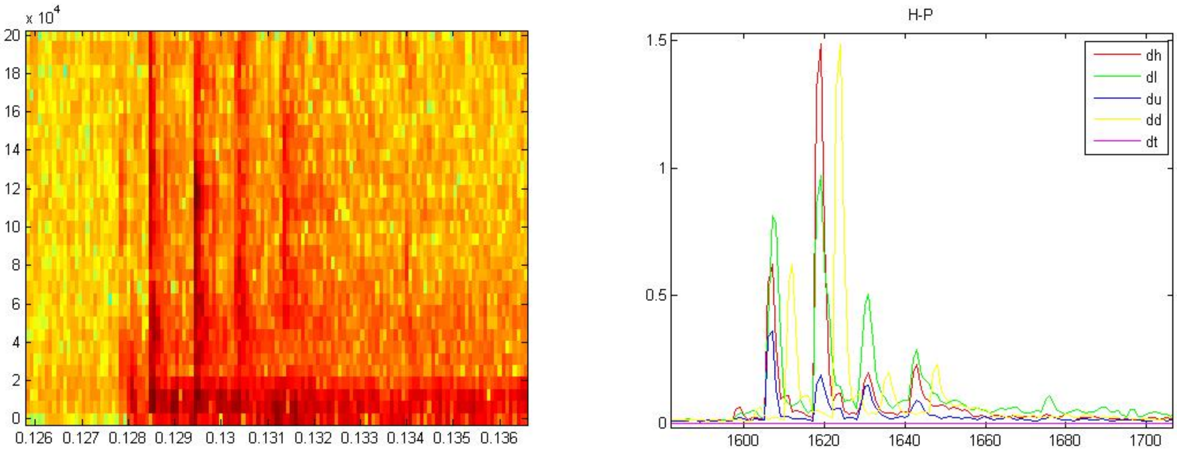


Figure 5. Processing crustacean click

These results show that SPUD is excellent at rejecting crustacean clicks and good at recognizing the clicks of harbour porpoise. The high false positive rate with bottlenose dolphin clicks will need further investigation but this level of false positives is acceptable as all of the false positives from the four datasets would have been rejected by the pulse train processing. Even the short sequence of regular clicks from the bottlenose dolphin would have been rejected because the pulse repetition rate was much slower than would be expected from a harbour porpoise.

It became clear from the testing that there were at least two types of crustacean click present in the test dataset. One has a comparatively low frequency content with peak energy around 2 kHz, while the other has a much more extended bandwidth with appreciable energy extending to over 100 kHz. Under some acoustic conditions, the spectrum of the high frequency clicks can be distorted such that a dip is introduced in the 60-90 kHz region and this leaves the higher frequency energy having very similar spectral characteristics to a harbour porpoise echolocation click. Because these false positives occur at random times they are rejected by the subsequent pulse train processing

The full implementation of the spectrogram correlation requires n*m 'multiply and add' operations. Part of the testing investigated whether this could be reduced without compromising performance. It was found that the column after the echolocation click often made no contribution to the performance because of the presence of multiple arrivals of the pulse and/or a reverberation tail. Removing these cells made no measurable difference to the performance of the processing. The columns before the pulse were reduced to a single column spaced three sample clocks ahead of the expected position of the click. This resulted in a small improvement in performance of the algorithm. Lastly the rows immediately above and below the expected position of the click were removed from the sum. This had no impact on the processing performance and an indeterminate impact on processing speed. The overall effect of refining the calculation was to approximately half the processing time while giving a small improvement in processing performance.

The pulse train processing testing showed that the algorithm is capable of following pulse trains. Unfortunately insufficient data was available to fully explore the capabilities of the current implementation. However, it became apparent that this processing becomes less effective as the number of echolocating animals increased. The same algorithm was tested with sperm whale clicks and it broke down completely when five or more animals were echolocating simultaneously. Further work is needed to make this algorithm more robust.

One problem encountered was that although it is possible to extract parameters from the pulse trains, such as minimum and maximum click rate and inter-pulse interval stability, there is no reference set of parameters for the harbour porpoise to use to test the measured parameters. There is also no information on how this reference set varies for different geographic areas. Before the pulse train processing can be fully proven it will be necessary to take a prototype SPUD system into the field and use it to build information on the pulse train statistics. Another problem is that the transmit beam is fairly narrow (~15 degrees [4]) and this results in short pulse trains as the beam illuminates the receiver while the animal is manoeuvring. It is only when the animal is close that long sequences can be built-up. Nevertheless, the statistical information from the combination of short busts of clicks should be sufficient to aid classification of the calls.

6 CONCLUSIONS

The work so far has demonstrated that the performance of the spectrogram correlator is acceptable but is not sufficient on its own to provide a reliable means of detecting and classifying harbour porpoise calls.

The addition of the pulse train processing does appear to reduce the false positive rate to acceptable levels. However this second stage of processing still has limited capability and should be able to provide improved performance, particularly when multiple echolocating animals are present.

Further testing of both parts of the processing is required before they can form a fully functional autonomous classifier of harbour porpoise echolocation clicks.

For pulse train processing to work successfully in autonomous mode there is a need for more detailed information on the statistics of the pulse sequences use by harbour porpoise. Similar statistics are also required for other animals which may be mis-classified by SPUD. This will require significant field effort at a number of field locations.

The pulse train processing to be used when multiple animals are present also needs to be reconsidered. A fragmented train processor may prove more robust in this situation.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the acoustic data sets proved by Mats Amundin, Paul Lepper and Ed Ceurstemont without which this work would not have been possible. The assistance provide by Peter Dobbins in making aspects of the testing possible is also gratefully acknowledged.

REFERENCES

1. Tregenza, N., Site acoustic monitoring for cetaceans - a self-contained sonar click detector, in: Proceedings of the seismic and marine mammals workshop London, 23-25 June 1998, M.L. Tasker and C. Weir, Editors. 1998, Sea Mammal Research Unit: St Andrews.
2. Mellinger, D.K. and C.W. Clark. A method for filtering bioacoustic transients by spectrogram image convolution. in OCEANS '93. 'Engineering in Harmony with Ocean*'. Proceedings*. 1993.
3. Mellinger, D.K. and C.W. Clark, Recognizing transient low-frequency whale sounds by spectrogram correlation. Journal of the Acoustical Society of America, 2000. **107**(6): p. 3518-3529.
4. Au, W.W.L., et al., Transmission beam pattern and echolocation signals of a harbor porpoise *(Phocoena phocoena).* Journal of the Acoustical Society of America, 1999. **106**(6): p. 3699-3705.
5. Teilmann, J., et al., Characteristics of echolocation signals used by a harbour porpoise (*Phocoena phocoena*) in a target detection experiment. Aquatic Mammals, 2002. **28**(3): p. 275-284.