Techniques for Segregation and Classification of Several Vocalizing Sperm Whales for AUV-based Localization Applications

Rajendar Bahl[1], Tamaki Ura[1], and Tetsuo Fukuchi[2]

[1] Institute of Industrial Science, The University of Tokyo, 4-6-1, Komaba, Meguro
Tokyo 153-8505, JAPAN

[2] System Giken Corp., 1436-1, Yahata, Chigasaki City
Kanagawa 253-0085, JAPAN

Abstract- **Sperm whales are amongst the most vocally active cetaceans. A fuller knowledge of these whales and their clicking behavior can be obtained by creating a picture of the relative locations of sperm whales in a region of interest and also by tracking them individually in three dimensions. This paper reports a major step in this direction by the development of automatic techniques for segregation of several clicking individuals and to generate a de-interleaved time series of their individual vocalizations. We discuss methods to extract the clicks from hydrophone data and for individual classification based on various click parameters: time-delay between hydrophones, inter-pulse-interval within a click, correlation with a reference, and the peak frequency. We present results of our techniques for automatic segregation and classification of several sperm whales using data collected from a 2-hydrophone array. These results demonstrate their suitability for a compact real-time implementation on AUV's.**

I. INTRODUCTION

Sperm whales (*Physeter macrocephalus*) are regarded as some of the most vocally active cetaceans. Both male and female individuals produce high intensity impulsive clicks at varying rates, depending on their activity. The clicks are categorized as "usual clicks", "slow clicks", "creaks" and "codas". Each click also has a multi-pulse structure. Various studies have been conducted [1-3] regarding the time and frequency characteristics of sperm whale clicks and creaks. Some insight has also been provided into the bio-sonar and inter-whale communication aspects of sperm whale clicks. However, conclusive evidence of the purpose of the clicks (and creaks) vis-à-vis the dive cycle has not been forthcoming so far [3].

A fuller knowledge of these whales and their clicking behavior can be obtained by creating a picture of the relative locations of sperm whales in a region of interest and also by tracking them in three dimensions. A major step in this direction would be to develop automatic techniques for real-time segregation of several clicking individuals and to generate a time series of their individual vocalizations. Section II discusses the signal processing techniques for extraction of sperm whale click parameters that are characteristic of individual animals. In section III we present results for automatic segregation and classification of several sperm whales. Section IV concludes and suggests a system for tracking individual diving sperm whales by an AUV.

II. PARAMETERS FOR IDENTIFICATION

A. *Data Collection*

The major known features of sperm whale clicks are:

- Click duration of 10-20 milliseconds
- Click rate usually between 0.5 to 2 clicks per second
- Creaks rates as high as 200 clicks per second
- Each click may have several pulses of decaying amplitude
- Bandwidth is very wide with significant energy up to 15 kHz.

The recordings of sperm whale vocalizations were made in early August 2002 during an expedition to Ogasawara islands. Fig. 1 shows the search trajectory and the locations (mainly at ocean depths between 1000-1500 meters) where sperm whale clicks were picked up.

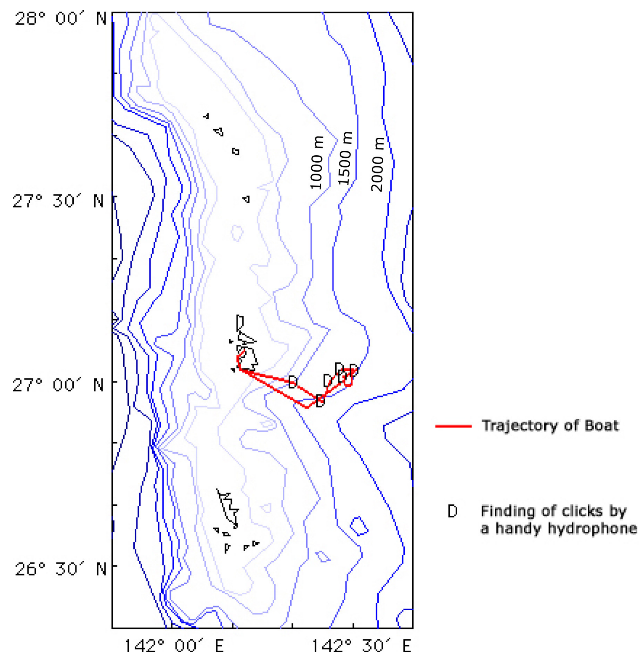

Fig.1. Map showing the search trajectory and locations where sperm whale clicks were picked up.

The sperm whale signals were picked up by a pair of hydrophones with sensitivity better than -200 dB re 1V/μPa (10 Hz to 30 kHz) separated by 75 mm. The hydrophone signals were passed through amplifiers with a gain of 40 dB ±0.1 dB between 1 kHz to 100 kHz, and recorded in 2-ch Sony TCD D10 DAT recorder at a sampling rate of 48 kHz. Recordings have been stored on CDROM in 16-bit stereo 'wav-file' format for further analysis. The major analysis platform is the MATLAB. The following paragraphs describe the step-by-step signal processing tasks for extraction of click parameters.

B. Signal Conditioning

The sperm whale sounds were recorded with hydrophones just below the sea surface. The recorded signals, therefore, have significant low frequency ambient noise (Fig. 2). This made it difficult to acquire the short clicks with reasonable S/N ratio.

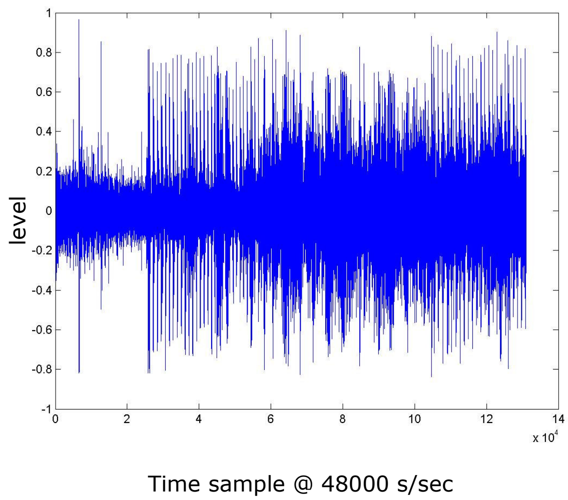

Fig. 2. Raw hydrophone signal.

It was decided to filter the recorded signals through a 5 kHz high pass linear phase FIR filter (31-sample Least Squares). Fig. 3 shows improvement in S/N offered by the filter since it rejects most of the low frequency interference. The filtered signals still contain significant click energy and have been found to retain the click parameters of interest. Subsequent signal processing has been performed using the filtered signals of both hydrophones.

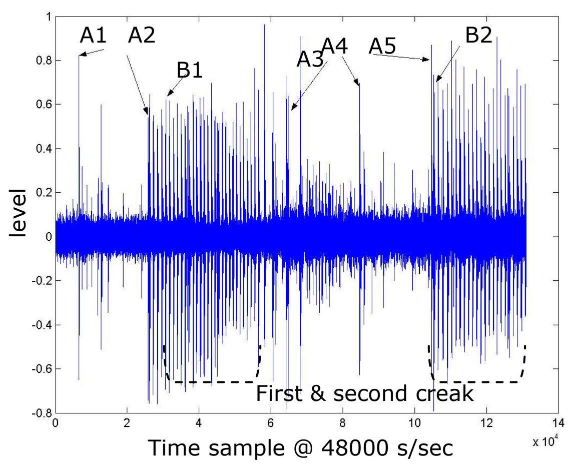

Fig. 3. High pass filtered signal.

C. Click Correlation

Correlation between clicks from the same whale is expected to be high, leading to whale identification. Thus, the first step is to cross-correlate an isolated click with a signal segment. The signal in Fig. 3 was cross-correlated by using certain clicks from the same sequence as references. For example, the correlation of the signal with the non-creak click A1 (Fig. 4) shows very low correlation with the creak clicks, except with itself and with click A2 just before the creak.

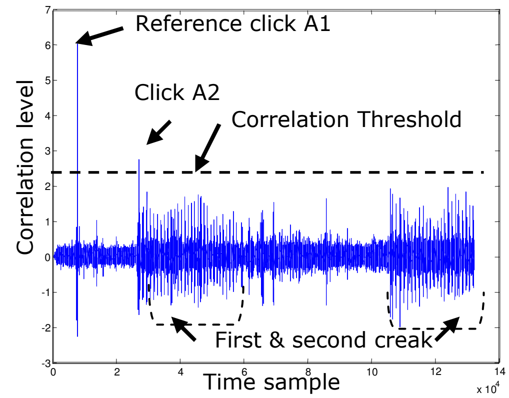

Fig. 4. Correlation with click A1.

Similarly, correlation of click B1 (Fig. 5) is very high (several clicks above the dotted line) with the neighboring clicks within the creak, while it is very low with click A1 and click A2.

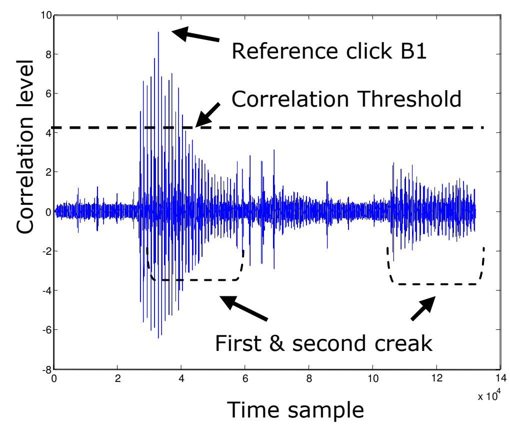

Fig. 5. Correlation with click B1.

The above results demonstrate that clicks within and outside the creak belong to different animals. However, it is also to be noted that the correlation falls with increasing distance from the reference click even for the same animal. We, therefore, need more robust analysis to associate clicks with specific individuals.

D. Inter-Pulse Interval

Another identifying characteristic of the clicking individual is the presence of multiple pulses within each click. The multiple pulses are thought to arise due to repeated reflections between the frontal and distal air sacs in the sperm whale's head. Since the distance between the two air sacs separated by the spermaceti oil sac depends on the size of the animal, it is expected that the time gap between multiple pulses would also depend on the size of the animal. The "inter-pulse interval" (IPI) may thus, be regarded as a distinguishing feature between individual animals. Fig. 6 shows a typical (filtered) click with the multiple pulse structure clearly visible. Cepstrum is an important tool that reveals the multiple-pulse structure of such signals. It is convenient to extract the IPI by simple thresholding of the cepstrum beyond the initial click duration.

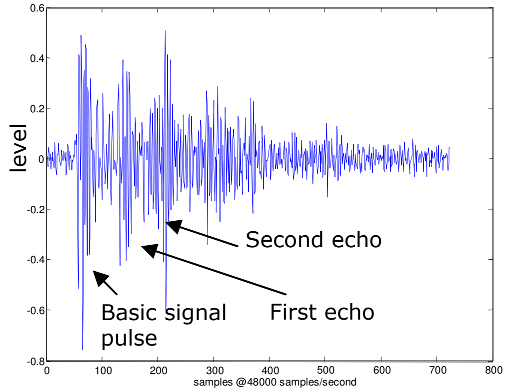

Fig. 6. Typical click showing multiple pulse structure.

Fig. 7 and Fig. 8 show the cepstrum and the IPI for two different clicks. It is evident from the cepstrum results that the two clicks arise from animals of different size as indicated by the difference in the IPI (1.6 msec and 1.5 msec): click B1 is perhaps generated by a smaller animal than that of click A1.

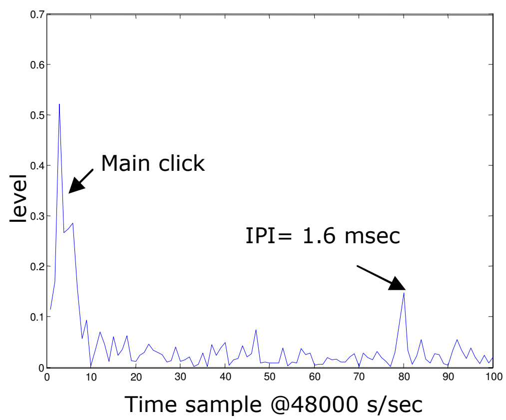

Fig. 7. Cepstrum of non-creak click A1.

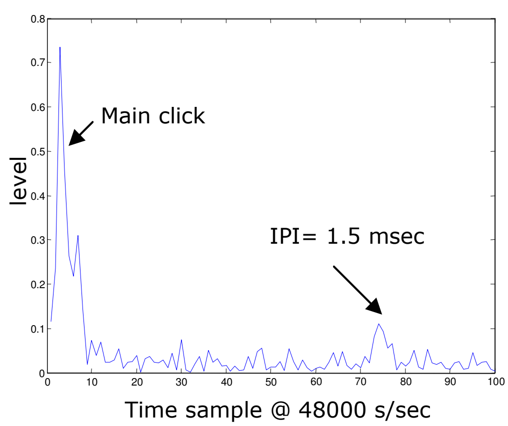

Fig. 8. Cepstrum of creak click B1.

We can, therefore, reinforce the results of correlation with cepstrum analysis to further discriminate the clicks and to distinguish the particular vocalizing animals.

E. Time Delay Estimation (TDE)

The above results have suggested a reasonable method to discriminate between clicks from the recording at a single hydrophone. There could, however, be ambiguities in associating the clicks with individual animals, unless spatial localization is also carried out using the time delay between the two hydrophones. We carried out TDE of isolated clicks between the two hydrophones (Figs. 9 and 10). Since we need very high time resolution to find the time delay between the two hydrophone signals, we performed interpolation by a factor of 16 to get an effective data rate of 48000x16 samples/sec.

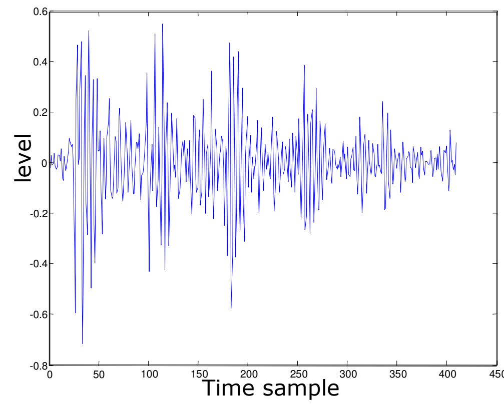

Fig. 9. Isolated click on left hydrophone.

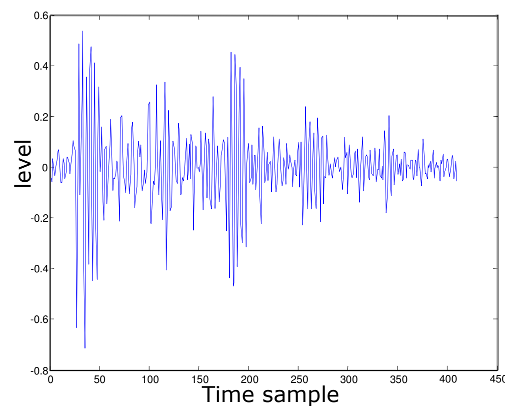

Fig. 10. Isolated click on right hydrophone.

Fig. 11 shows the auto-correlation peak of the left hydrophone signal occurring at time sample #245.

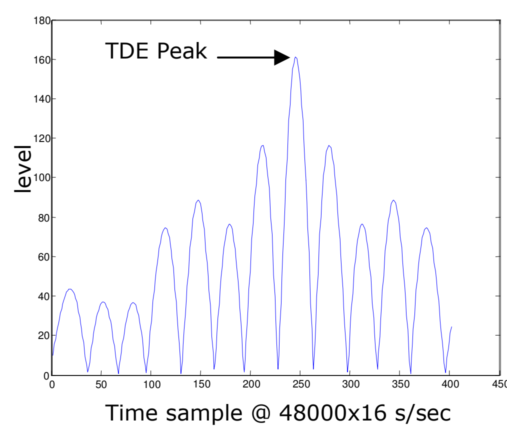

Fig. 11. Expanded view of auto-correlation peak of left hydrophone.

Fig. 12 shows the cross-correlation peak between the two hydrophones at sample #225. The time difference of 20 samples is an indicator of the angle of arrival of the wavefront at the hydrophones.

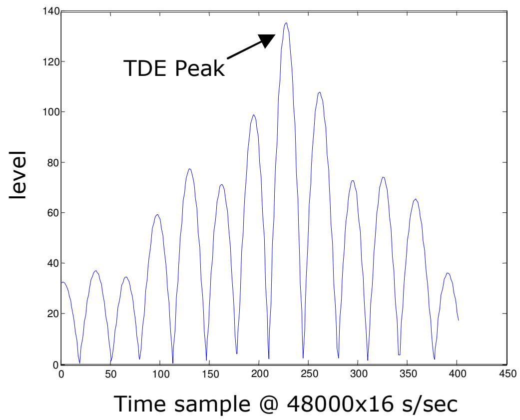

Fig. 12. Expanded view of the cross-correlation.

F. Preliminary Identification Results

We applied the above TDE technique on various clicks indicated in Fig. 3. It was confirmed that the set of clicks A1, A2, A3, A4 and A5 have similar time delay and so they arrive from one location. Clicks B1 and B2 from the two creaks have similar time delay but are different from the previous set, hence they come from another location. In view of the correlation results, IPI values and TDE estimates obtained from the recorded data, it can now be broadly said that clicks A1-A5 come from one animal (whale A), while clicks B1 and B2 (in the two creaks) belong to another animal (whale B). The preliminary identification results for the two specific whales are indicated in Fig. 13.

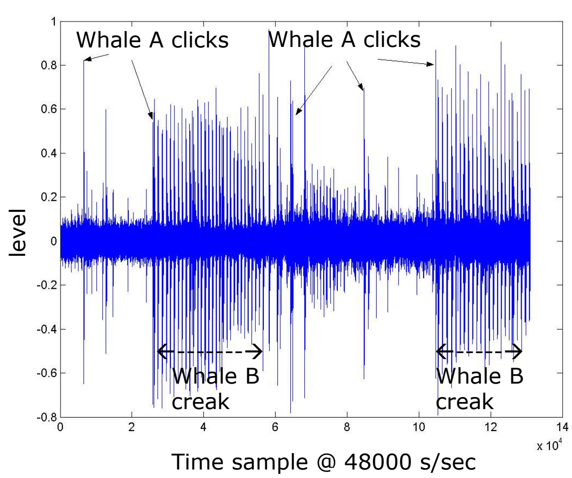

Fig. 13. Preliminary identification results.

A more comprehensive analysis based on the above techniques would reveal the identity of several other clicks that are present in the data. The above analyses on sperm whale clicks provide us a basis for characterization of the individual animals. Correlation is useful in segregating adjacent clicks, but it is not a strong enough indicator beyond a few clicks even for the same animal. IPI analysis helps in determining animal size and can be used in association with the correlation level for identifying the concerned animal. Finally, TDE between two hydrophones provides us with spatial evidence for segregating animals. The three parameters, when used together are seen to help in identifying the individual sperm whales. In the following, we use the above parameters and also the spectrum peak of the click for automatic real-time classification of several whales.

III. AUTOMATIC REAL-TIME CLASSIFICATION

A. Click Extraction

A basic task for real-time processing is to automatically extract clicks as they occur in the hydrophone record. Clicks from an individual whale may be as close as 5 milliseconds during “creaks”. On the other hand, click duration may be as large as 10-20 msec. As seen in the data, there may also be several independently clicking whales in a region (Fig. 14). As a result, the signals picked up by a hydrophone array would consist of overlapping and interleaved clicks of several sperm whales.

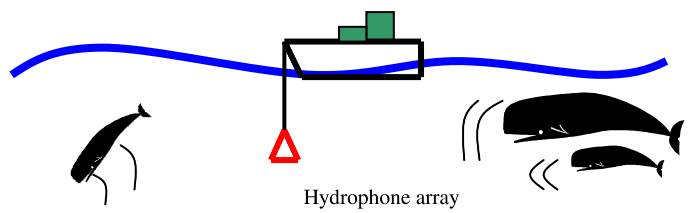

Fig. 14. Multiple vocalizing sperm whales.

Fig. 15 shows a 20 second record of filtered data from one of the hydrophones in a typical multiple whale scenario.

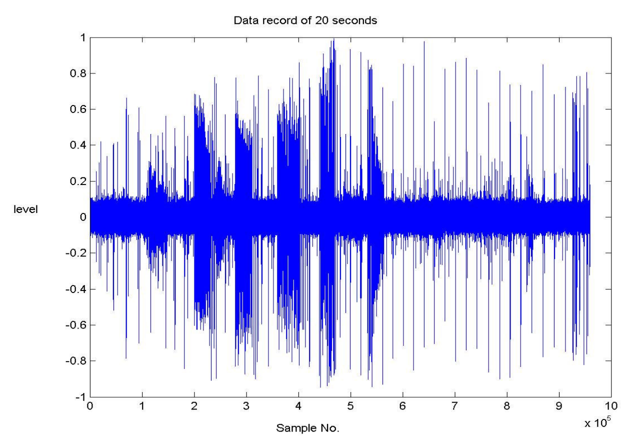

Fig. 15. Data record of 20 seconds.

It is, therefore, considered necessary to have a time resolution of 5 msec for click extraction. The data from a reference hydrophone needs to be analyzed in contiguous 5 msec segments. We extract valid clicks by computing energy over each 5 msec segment and comparing it with a threshold, while noise spikes having small energy shall be rejected by the threshold. The threshold can be made adaptive to cater for slow variations in ambient noise. We have found that for our data, a 10-20 second average of the segment energy is adequate to estimate the background noise energy. A threshold of k-times the average (e.g., k=1.5) is used for detecting the presence of a click.

Once the presence of a click is detected, the click is extracted based on the fact that its multiple pulse structure typically lasts for about 10 msec. The click data is extracted using the detected segment and those preceding and succeeding it in a manner to capture the leading portion of the click. Fig. 16 shows the click extraction procedure.

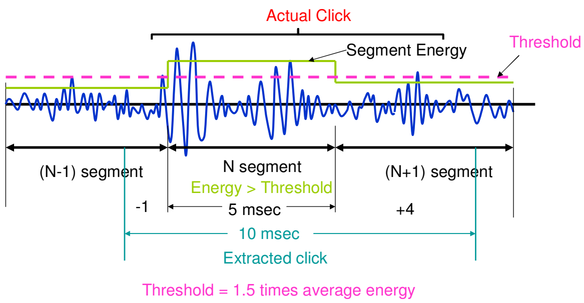

Fig. 16. Click detection and extraction.

The energy level of the detected and extracted clicks on 20 seconds of data of Fig. 15 is shown in Fig. 17. All the extracted clicks are likely to be those of sperm whales in the same region.

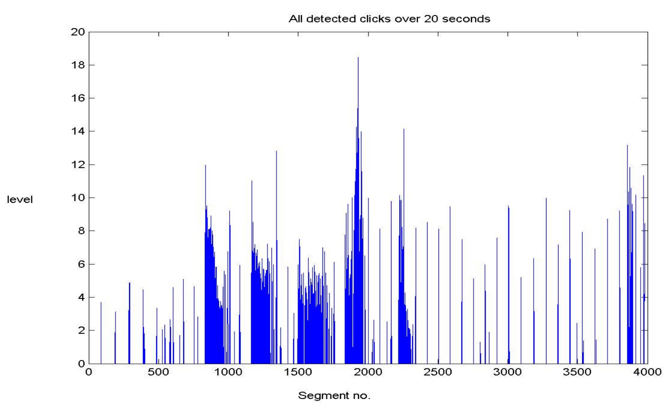

Fig. 17. Extracted clicks for 20 second data.

B. Basic Classification Strategy

The parameters used for classifying sperm whales from their extracted clicks are:

- TDE: time delay between two hydrophones used as a primary parameter,
- CORR: normalized correlation peak with a particular reference,
- IPI: Inter-Pulse-Interval, and
- SPEC: peak frequency of click spectrum.

The classification procedure takes into account the fluctuations expected in the above parameters due to environmental and biological reasons: ambient noise, multipath and changes in click generation mechanism. The allowable range of parameter values are: TDE within ± 0.15 points (± 3 microsec), CORR within 40%, IPI within ± 8 samples (± 0.17 msec), and SPEC within ± 8 points (± 800 Hz).

The key feature of classification is the updating of reference click with the most recently classified click. The strategy adopted is as follows:

1) Start with an initial reference click.
2) For the test click, check if TDE is within limits, else select next reference.
3) Then check if any one of CORR, IPI, SPEC are within limits, else select next reference.
4) If a match is found, then label new click with the reference class, and also update the reference click with the new click provided there is a separation of more than 2 segments from the previous match (i.e., >10 msec).
5) If no match is found, then label the click as a new class and add to the reference clicks list.

The classification procedure was tested on the clicks extracted from the data record that apparently has several “clicking and creaking” whales (Figs. 15 and 17). The result of the basic classification is shown as a histogram in Fig. 18. The procedure recovered 3 significant classes out of 34 possible classes.

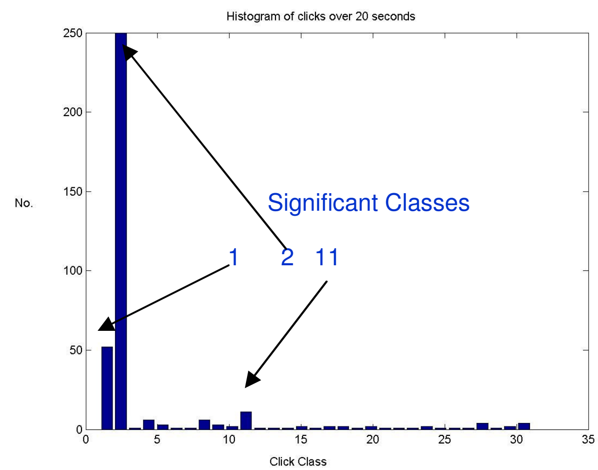

Fig. 18. Histogram of basic classification.

The clicks of these 3 significant classes are de-interleaved on the basis of the classification. They are shown in terms of energy levels in Figs 19-21.

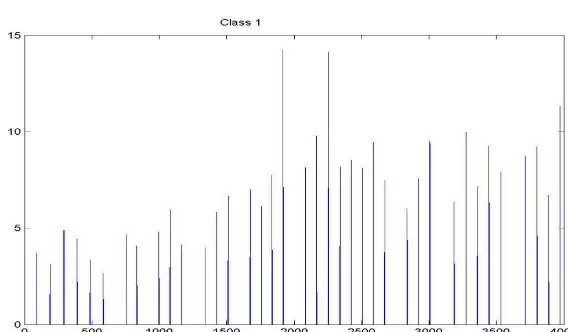

Fig. 19. Class 1 clicks (whale A).

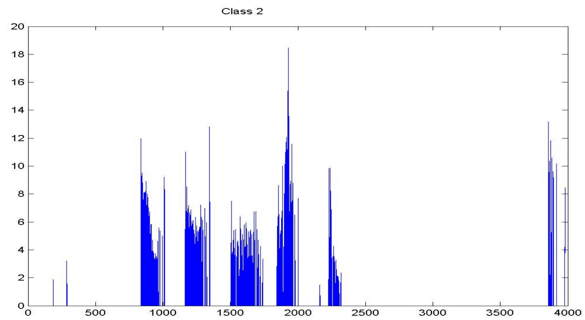

Fig. 20. Class 2 clicks (whale B).

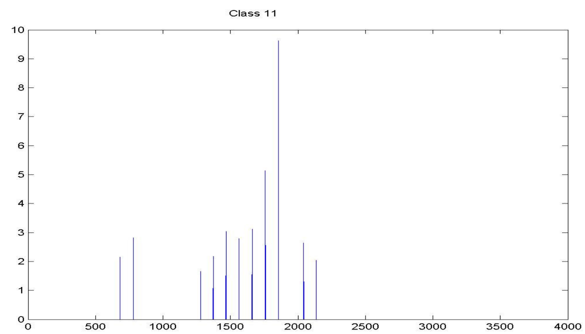

Fig. 21. Class 11 clicks (whale C).

C. Refinements

Though the classification technique has recognized the dominant presence of 3 whales, there are some notable shortcomings in the basic strategy described above which causes ambiguities in classification. The first problem is that it does not account for the possibility that IPI can sometimes be twice the usual value due to peculiarities of click generation, propagation and cepstrum processing. This is most evident in the extracted IPI of the creaking whale B (class 2) as shown in Figure 22.

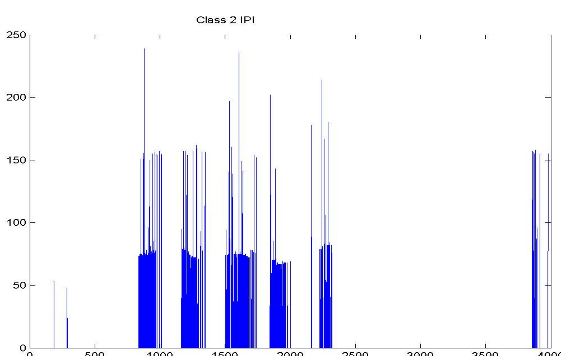

Fig. 22. IPI variation of whale B.

Interestingly, out of a total of 51 clicks classified as Class 1, almost 40% of the updated clicks were recognized on the basis of IPI alone. It is, therefore, important that IPI be given utmost care in its use as a classification parameter.

The second refinement concerns the possibility of misclassification of a class as a new class. A particular click may get classified as a new class at some point instead of an existing class, and subsequently match with the previously existing class again. This class "forking" and "merging" needs to be curtailed, if not entirely preventable.

The modified technique, therefore, permits IPI values to be double of the class IPI, and in case a click matches with more than one class, then the class having more clicks is retained in favor of the less frequent class. Histogram resulting from the modified technique is presented in Fig. 23.

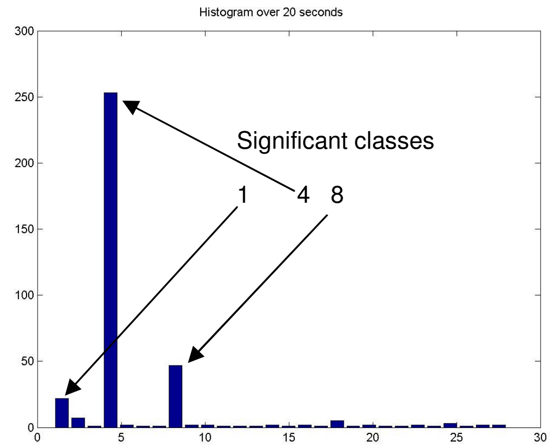

Fig. 23. Histogram for modified technique.

Comparison with the histogram of Fig. 18 shows a better concentration of clicks into 3 major classes (renumbered). The number of classes has also reduced to 28 from 31. It can be said that most of the other classes belong to whales whose clicks are inconsistently received in the 20 second record.

The time-domain de-interleaved results of the refined classification are shown in Figs. 24-26 in terms of energy levels.

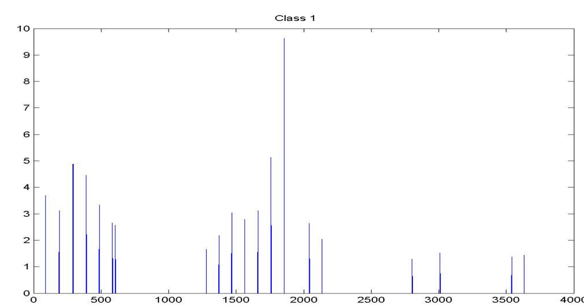

Fig. 24. Class 1 clicks (whale C).

It is noted that the initial clicks of whale C that were previously misclassified as whale A (in Fig. 19) have now been properly classified as whale C (initial clicks in Fig. 24).

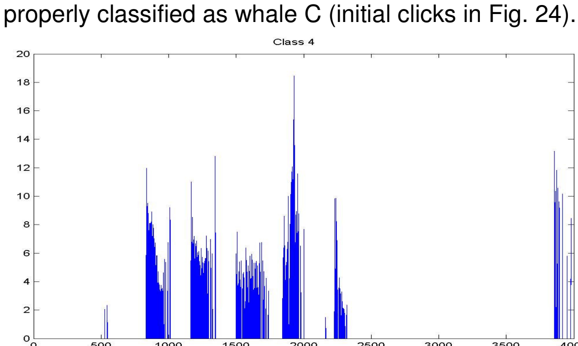

Fig. 25. Class 4 clicks (whale B).

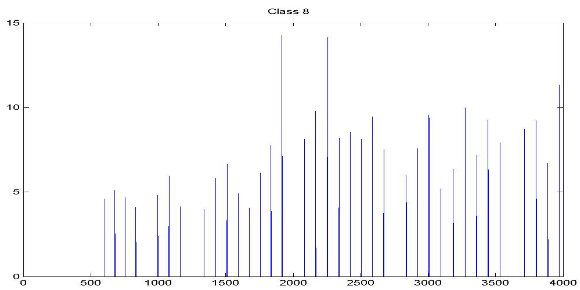

Fig. 26. Class 8 clicks (whale A).

IPI is an important whale-specific parameter for the three classes. Our case study showed that class 1 (whale C) has a dominant IPI of 50 units, class 4 (whale B) has IPI of 75 and class 8 (whale A) has IPI of 80. Hence, the three whales are expected to be of different sizes. It can also be seen that the three whales have a different click pattern perhaps due to different activities of the individuals.

IV. CONCLUSION AND FUTURE PLANS

The automatic classification technique has been demonstrated to provide good results for segregation of several vocalizing whales in real-time. No attempt has been made to incorporate features based on time-history as that will make the system non-real time for tracking applications that are subsequently envisaged. Further work is planned to test the technique in actual sea-trials. A schematic of the proposed system is shown in Fig. 27.

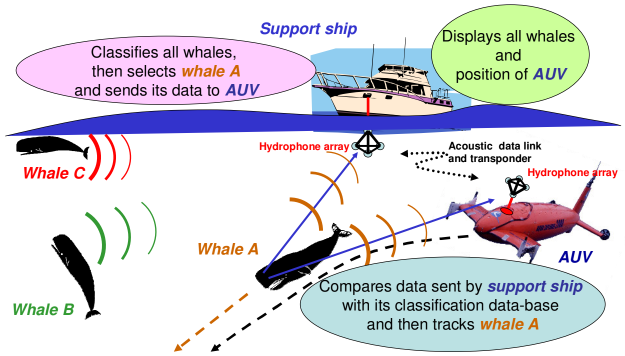

Fig. 27. Proposed system for tracking selected sperm whale with an AUV.

The AUV will work in conjunction with a surface support ship. Both vehicles will develop and maintain a data-base of the various clicking sperm whales and keep track of each other's location with help of GPS, depth sensors, SBL arrays and acoustic transponders. The identified whale will then be tracked by the AUV, while the support vessel displays the locations of the AUV and the tracked whale. A set of experiments is planned to develop the sperm whale tracking system on these lines in the near future.

ACKNOWLEDGMENTS

The authors wish to express their sincere appreciation of Dr. Kyoichi Mori of Ogasawara Whale Watching Association for his able guidance in locating the sperm whales. They also thank Dr. Tomonari Akamatsu of National Research Institute of Fisheries Engineering for his advice on sperm whale vocalizations. The assistance provided by Mr. Junya Ura of Waseda University, Mr. Ryoji Shimizu and Mr. Yuichi Shirasaki in recording the data is gratefully acknowledged.

This work was supported by the "grants-in-aid for scientific research" program of JSPS.

REFERENCES

[1] J.C. Goold and Sarah E. Jones, "Time and frequency domain characteristics of sperm whale clicks," *J. Acoust. Soc. Am.,* Vol. 98, pp 1279-1291, 1995.

[2] G. Pavan, T.J. Hayward, J.F. Borsani, M. Priano, M. Manghi, C. Fossati, and G. Gordon, "Time patterns of sperm whale codas recorded in the Mediterranean Sea 1985-1996," *J. Acoust. Soc. Am.,* Vol. 107, pp 3487-3495, 2000.

[3] N. Jacquet, S. Dawson, and L. Douglas, "Vocal behavior of male sperm whales: Why do they click?," *J. Acoust. Soc. Am.,* Vol. 109, pp 2254-2259, 2001.