研 究 速 報

Automatic Real-time Segregation and Classification of Multiple Vocalizing Sperm Whales

Rajendar BAHL* and Tamaki URA*

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". While there is no consensus regarding the exact purpose of the various types of clicks, there is a general agreement that the clicks are used for echo-location and some type of inter-whale communication.

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. This paper is a follow-up of our previous paper[1] that discussed signal processing techniques for extraction of sperm whale click parameters. We now present results of our techniques for automatic segregation and classification of several sperm whales. The techniques are subjected to data collected from a pair of hydrophones as reported in[1] .

II. CLICK EXTRACTION

A basic task for real-time processing is to automatically extract clicks present in the hydrophone data. 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. There may also be several independently clicking whales in a region (Fig. 1). As a result, the signals picked up by a hydrophone array would consist of overlapping and interleaved clicks of several sperm whales. Fig. 2 shows a 20 second record of data on one of the hydrophones in a multiple whale scenario.

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. An energy-based threshold is expected to perform better than a simple level-based threshold since the clicks have significant energy compared to transient noise spikes. 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

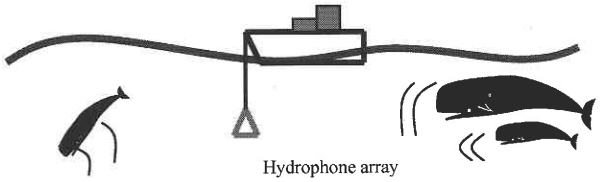

Fig. 1 Multiple vocalizing sperm whales

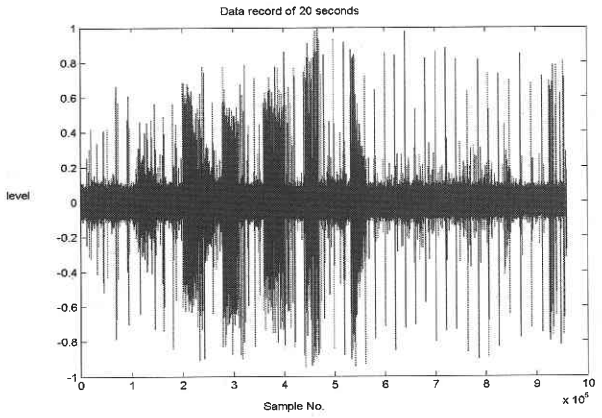

Fig. 2 Data record of 20 seconds

*Underwater Technology Research Center, IIS, University of Tokyo

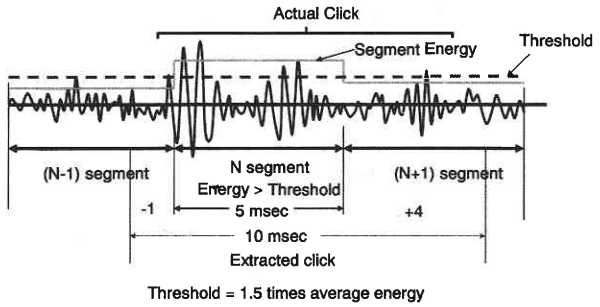

Fig. 3 Click detection and extraction

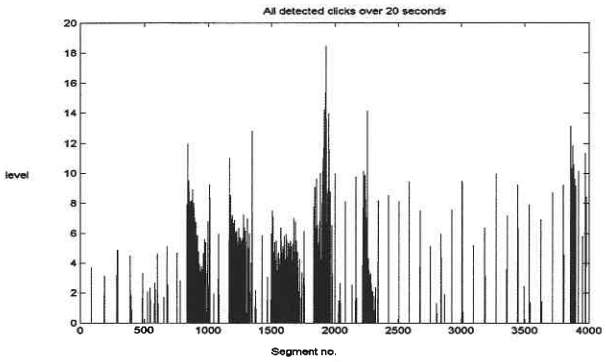

Fig. 4 Extracted clicks for 20 second data

our data, a 10–20 second average of the segment energy is adequate to find 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 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. 3 shows the click extraction procedure.

The result of click extraction on 20 seconds of data of Fig. 2 is shown in Fig. 4. All the extracted clicks are likely to be those of sperm whales in the same region.

III. BASIC CLASSIFICATION

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

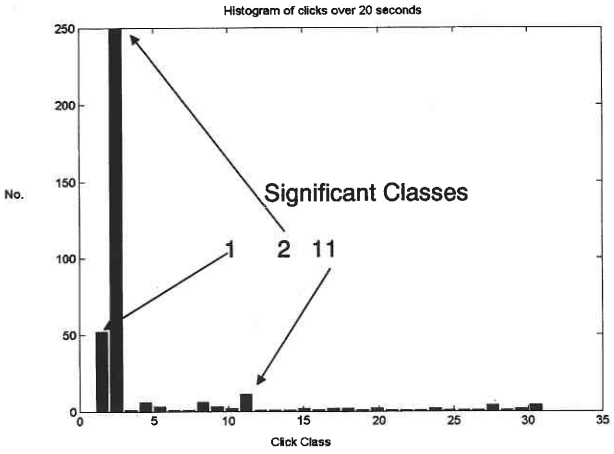

Fig. 5 Histogram of basic classification

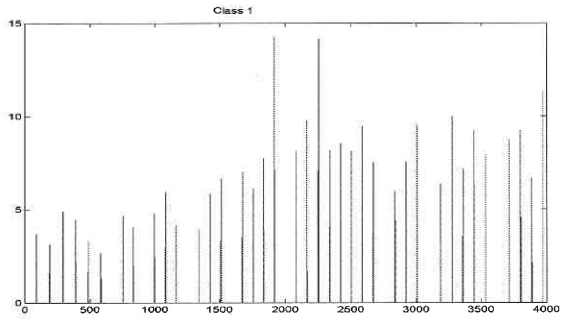

Fig. 6 Class 1 clicks (whale A)

- 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:

A) Start with an initial reference click.

B) For the test click, check if TDE is within limits, else select next reference.

C) Then check if any one of CORR, IPI, SPEC are within limits, else select next reference.

D) 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

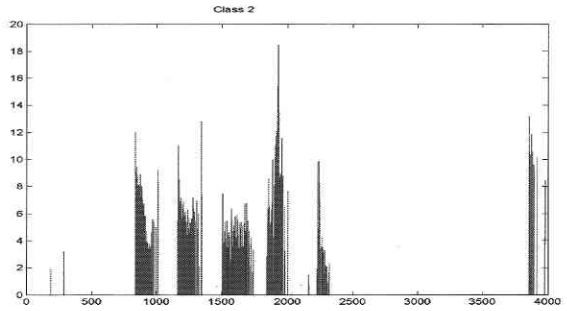

Fig. 7 Class 2 clicks (whale B)

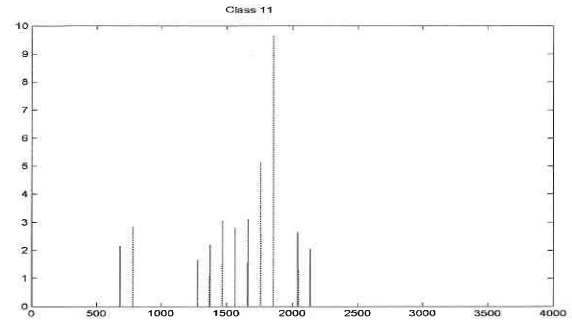

Fig. 8 Class 11 clicks (whale C)

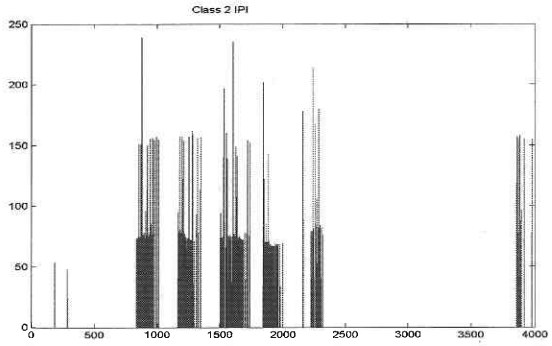

Fig. 9 IPI variation of whale B

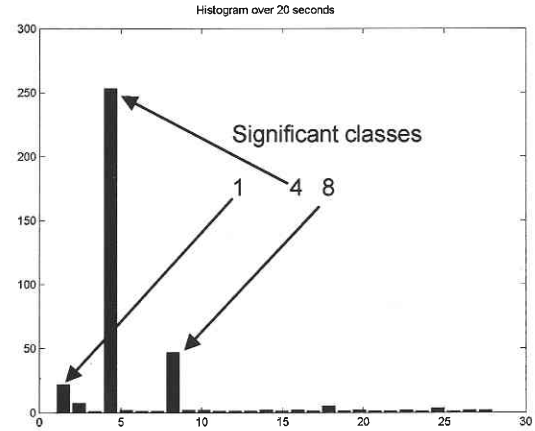

Fig. 10 Histogram for modified technique

the previous match (i.e., > 10 msec).

E) If no match is found, then label the click as a new class.

The classification procedure was tested on the clicks extracted from the data record that apparently has several "clicking and creaking" whales (Figs. 2 and 4). The result of the basic classification is shown as a histogram in Fig. 5. The procedure recovered 3 significant classes out of 31 possible classes. The de-interleaved clicks of these 3 classes are shown in Figs 6–8.

IV. REFINEMENTS

Though the technique has recognized the dominant presence of 3 whales, there are some notable shortcomings in the basic classification procedure 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. This is most evident in the IPI of the creaking whale B (class 2) as shown in Figure 9.

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 somehow get classified as a new class at some point instead of an existing class, and subsequently match with the previous class again. This class "forking" 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. 10.

Comparison with histogram of Fig. 5 shows a better concentration of clicks into 3 major classes. 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 inconsistent in the 20 second record.

The time-domain de-interleaved results of classification are shown in Figs. 11–13 in terms of energy levels. It is noted that the initial clicks of whale C that were previously misclassified as whale A have now been properly classified as whale C.

IPI is an important whale-specific parameter for the three

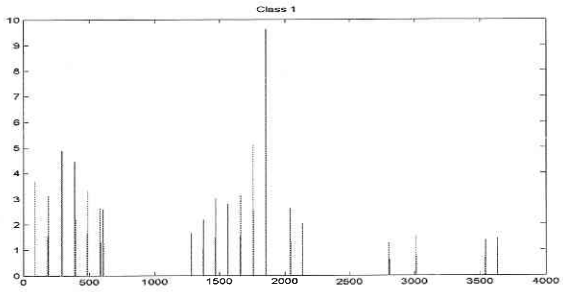

Fig. 11 Class 1 clicks (whale C)

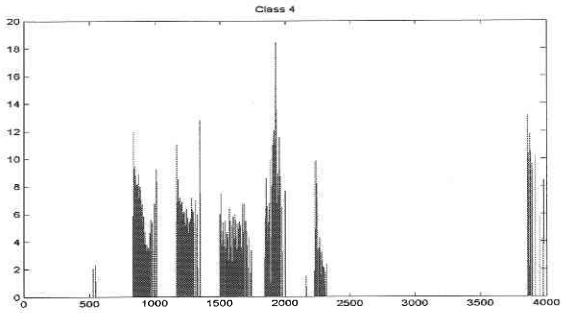

Fig. 12 Class 4 clicks (whale B)

classes: 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 size. It can also be seen that the three whales have a different dominant inter-

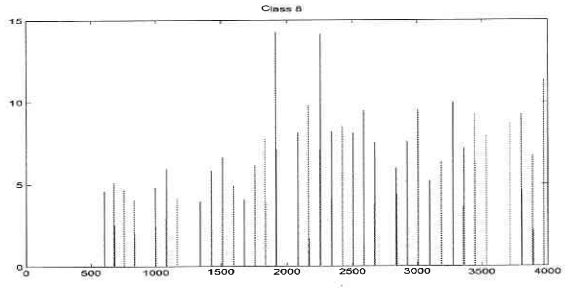

Fig. 13 Class 8 clicks (whale A)

click interval perhaps due to different activities.

IV. CONCLUSIONS

The automatic classification technique has been demonstrated to provide good results for segregation of several vocalizing whales in real-time. No effort has been made to improve classification based on time-history as that will make the system non-real time. Further work is planned to test the technique in actual sea-trials later this year.

(Manuscript received, FEBRUARY 20, 2003)

V. REFERENCES

1) Bahl, R., Ura, T., and Fukuchi, T. (2002). "Towards identification of sperm whales from their vocalizations," Seisan-Kenkyu Bimonthly Journal of IIS, University of Tokyo, Vol. 54, No. 6, 51–55.