論文の内容の要旨

Abstract of Dissertation

Acoustic Classification of Whales and Dolphins
(音響による鯨類の個体識別)

マズハー　スレーマン

(Mazhar Suleman)

(本文) Abstract

This research presents an attempt to build an automatic recognition system for individual humpback whales and bottlenose dolphins based on their audio records. Such recognition systems are an improvement upon bioacoustics-based tracking methods for population census of endangered species for two reasons. Firstly, when coupled with tracking, they provide a unique automated population assessment tool which is free of human biases and allows large area coverage during ecological surveys conducted for conservation of endangered species. Secondly, these methods provide a more robust substitute for inconvenient and error-prone photo-ID tools used for monitoring migratory patterns of species like humpback whales.

Although both species have a large repertoire of sounds [2,3] which serve a number of unique functions, this work targets the humpback whales songs and the clicks of bottlenose dolphins to accomplish this task. Although both type of vocalizations, obtained from unrestricted individuals and recorded using passive systems in our case, pose a challenging classification problem in a dynamically changing environment yet this selection is made because the frequency of occurrence for clicks and songs is higher than any other type of vocalizations in respective species.

Classification of humpback whale

The architecture of humpback classification system is shown in Fig. 2(a). Audio data, recorded using a single hydrophone (shown in Fig. 1), is pre-processed into smaller file-segments and down-sampled at 8 KHz for further processing. This is followed by song unit detection using bar crossing rate (BCR) algorithm [4,5] . Cepstrum feature vectors are extracted from each song unit as explained in Fig. 2 (b). The feature vectors coming from individual humpback whales are separated and used to train the support vector machine (SVM) model, to be used later for classification. SVM was used because of its accepted optimal performance in literature [1] (Fig. 3). During classification stage, feature vectors in test set are grouped together to make utterance vectors. The final classification decision for assigning any utterance vector to a particular whale class is taken by following three threshold rules.

$$\text{Class of } X_k = \begin{cases} \arg \max\limits_{1 \le i \le D} P_i \\ \arg \max\limits_{1 \le i \le D} P_i > 30\ \% \\ \arg \max\limits_{1 \le i \le D} P_i > 50\ \% \end{cases} \tag{1}$$

, where D is the test utterance vector length and P_i represents the decision taken by multi-class SVM model, based on one-against-one voting approach, for single feature vector from test set.

For humpback whales, we have tested our algorithms on both clean and noisy data [4,6]. Moreover, the accuracy of the system is tested for data sets of different individuals, recorded in different years [5]. The suggested system is robust enough to classify whales with accuracy greater than 90%. It is observed that a test utterance vector corresponding to two minute duration is sufficient to achieve this classification accuracy for good quality songs [4]. The noisy conditions may impact the accuracy but use of suitable filtering and spectral subtraction, based on silence parts of recording, shows remarkable improvement in accuracy results. The only problem is the classification of animal records with a gap of more than a few years [6]. In that case, the system may not be able to perform best. But still accuracy remains above 50%. Similarly, using pooling and threshold-based decision methodology, provision is provided to incorporate arrival of new animals.

Classification of bottlenose dolphin

The architecture for bottlenose dolphin tracking and classification system is shown in Fig. 5. The acoustic data was recorded from a 5-hydrophone array as shown in Fig. 4 (a). The data was collected in a target-touch experiment in which two dolphins were trained to approach and touch a target of 2 MHz hydrophone place at a distance of 2 m from the 5-hydrophone array. Fig. 4(b) shows the scenario of target-touch experiment with a subject dolphin approaching the target (2 MHz hydrophone), held by trainer. The direct and reflected clicks, above a user-specified energy threshold, are successfully detected using a running energy sum algorithm. Time and path delays are approximated with an improved correlation-based method applied to 5-channel data. Target position was localized using following formulae for range r, horizontal angle θ and elevation angle ϕ

$$r = \frac{2B - d_1^2 - d_3^2}{2(d_1 + d_3)}$$

$$\theta = \tan^{-1}\left(\frac{b^2 - d_4^2 - 2rd_4}{b^2 - d_5^2 - 2rd_5}\right)$$

$$\phi = \cos^{-1}\left(\frac{d_1^2 - d_3^2 + 2r(d_1 - d_3)}{4rB}\right) \qquad (2)$$

Short-duration classification of different dolphins is performed using track-association of each unlabelled click using three approaches which considered position information, click characteristics (similarity pattern in time domain) and hybrid features (position/click characteristics) respectively. Comparison suggests superiority of hybrid feature technique over two former techniques. For track-independent classification, a number of features such as spectral coefficient, cepstrum coefficients, temporal click patterns, beam patterns of each click and corresponding equivalent aperture size for dolphin transmission system have been analyzed.

Classification problem of bottlenose dolphin clicks is especially challenging because of highly directional beam patterns which results in different click characteristics (energy, temporal click shape, dominant frequencies) at different orientation [7]. Even with multi-hydrophone system, such as one described above, established conventional algorithms for signal detection and delay estimation cannot be trusted due to complexities resulted by click reflections in shallow depths and narrow beam pattern of the clicks. Therefore a semi-autonomous analysis was adopted wherein derived click parameters (time delays, position coordinates and their one and two dimensional projections in space, dominant frequencies on three hydrophones along vertical baseline etc) could be extracted automatically and later examined using graphical user interface (GUI) as shown in Fig. 6. The GUI has options to change or re-calculate some of these features. Moreover new clicks can be added or automatically identified clicks can be removed if desired. Provision is provided for detection of low-energy clicks (earlier missed by automatic method) in manual as well as in semi-autonomous mode where a user is allowed to provide a reference click and a relative energy threshold. For bottlenose dolphin, the system ensures reliable observation and identification in short-duration. However, identification over long periods is still open for further exploration and research.

Overall, both systems are expected to be useful for recording valuable information such as cetacean movement, behaviour and their population density estimation in a particular region and provide a reliable alternative for currently prevalent techniques in the field for cetacean conservation.

References

[1] L. Bottou and C. J. Lin. "Support Vector Machine Solvers", Large Scale Kernel Machines, L. Bottou, O. Chapelle, D. DeCoste, and J. Weston editors, MIT Press, Cambridge, MA, pp. 1–28, 2007.
[2] A. K. Stimpert and W. W. L. Au., M. P. Johnson, D. N. Wiley, "Megapclicks: Tagged humpback whales produce click trains and buzzes during foraging" (A) *Journal of the Acoustical Society of America*, Vol. 120, No. 5, pp. 3012-3012, Nov. 2006.
[3] W. W. L. Au, "Hearing in whales and dolphins: An overview", in W. W. L. Au and A. N. Popper and R. R. Fay, Hearing by whales and dolphins pp. 1-42, 2000.
[4] S. Mazhar, T. Ura and R. Bahl, "Vocalization based individual classification of humpback whales using support vector machine", *IEEE/OES Oceans' 2007, Vancouver*, 2007.
[5] S. Mazhar, T. Ura and R. Bahl, "Effect of Temporal Evolution of Songs on Cepstrum-based Voice Signature in Humpback Whales", *IEEE/OES Oceans' 2008, Kobe*, 2008.
[6] S. Mazhar, T. Ura and R. Bahl, "An analysis of Humpback whale songs for individual classification" (A) *The Journal of the Acoustical Society of America*, Vol. 123, No. 5, p. 3774, 2008.
[7] W. W. L. Au, The Sonar of Dolphins, Springer-Verlag, New York, 1993.

Fig. 1: System used for humpback whale song recording

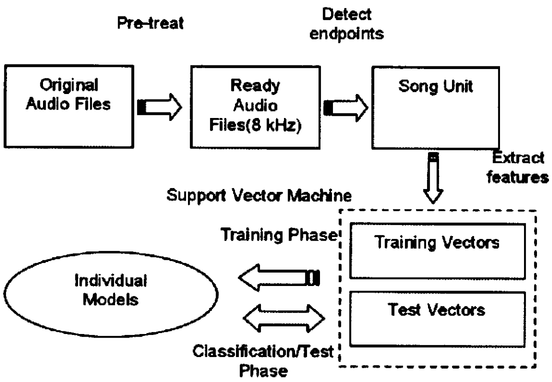

Fig. 2 (a): Algorithmic diagram for humpback whale classification

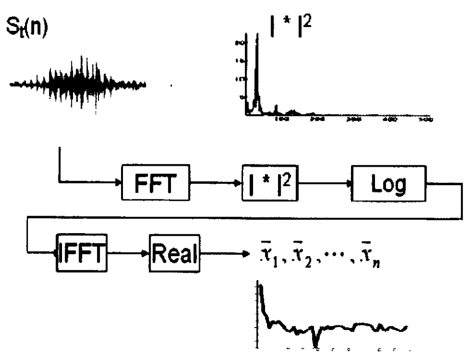

Fig.2 (b): Extraction of cepstrum coefficients

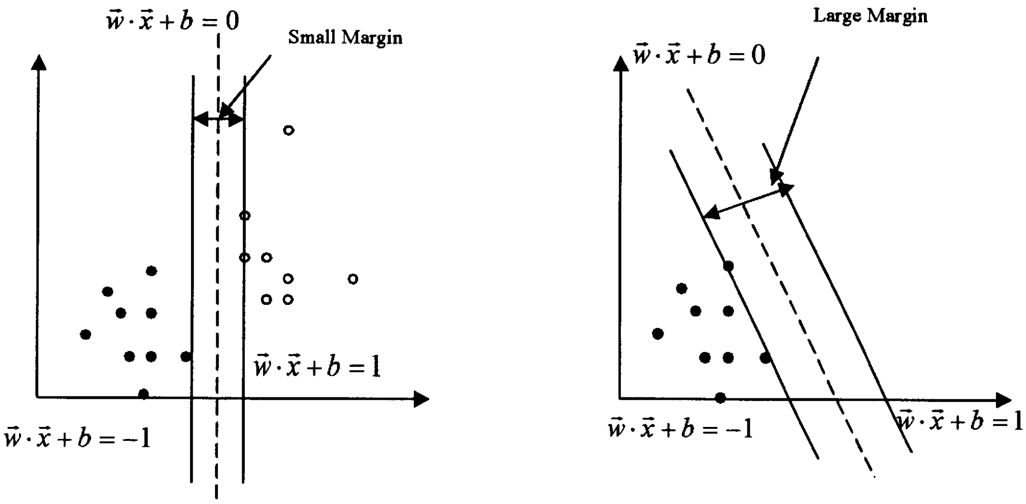

Fig. 3: Support vector machine used for optimal classification ensures selection of large margin classifier

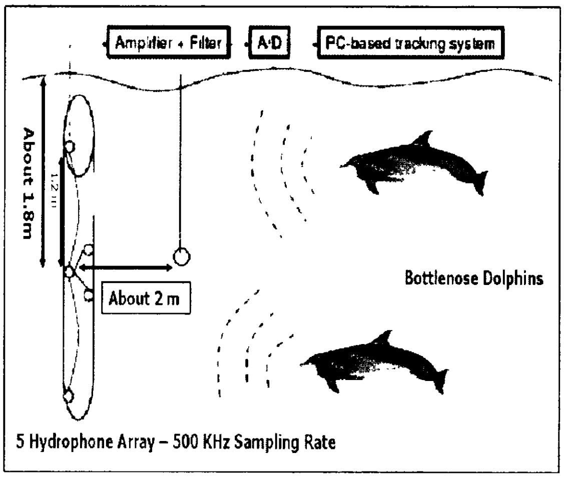

Fig. 4 (a): 5-Hydrophone system for bottlenose classification

Fig. 4 (b): 5-A scenario of target-touch experiment

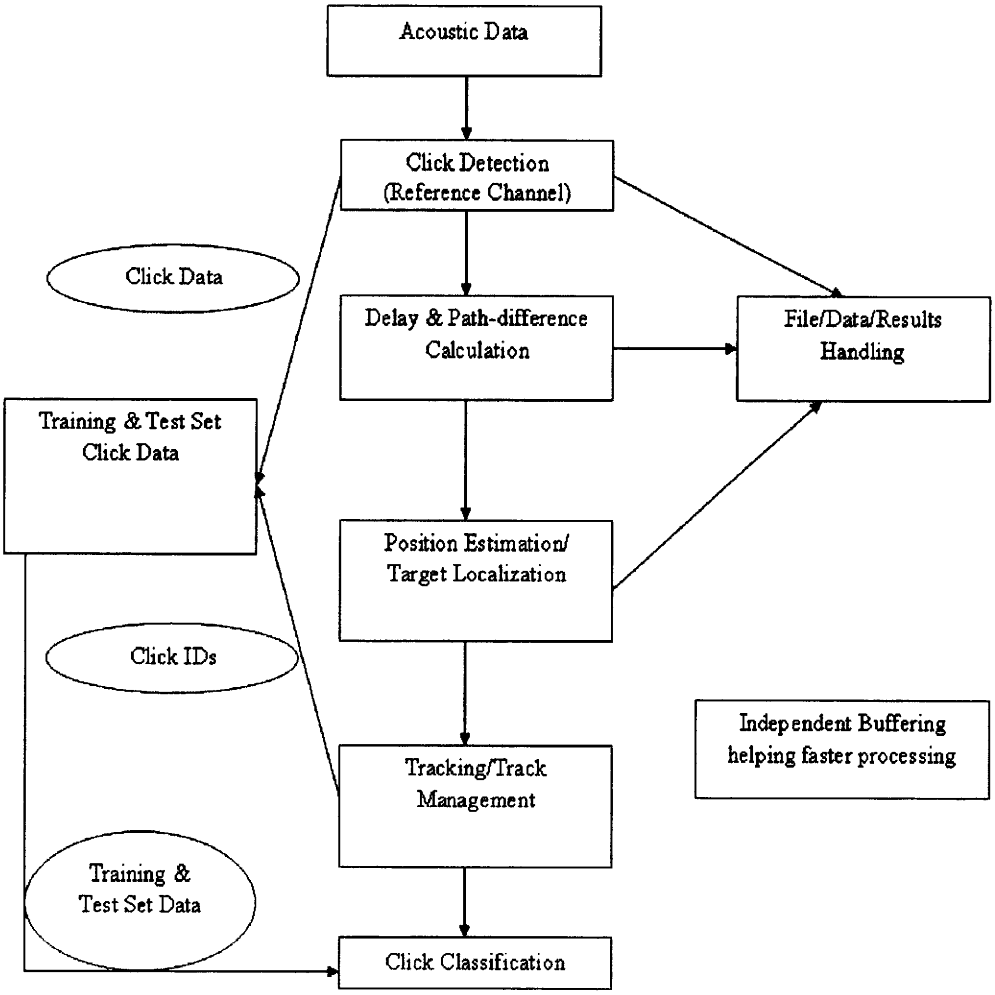

Fig. 5: Algorithmic block diagram for bottlenose tracking and classification

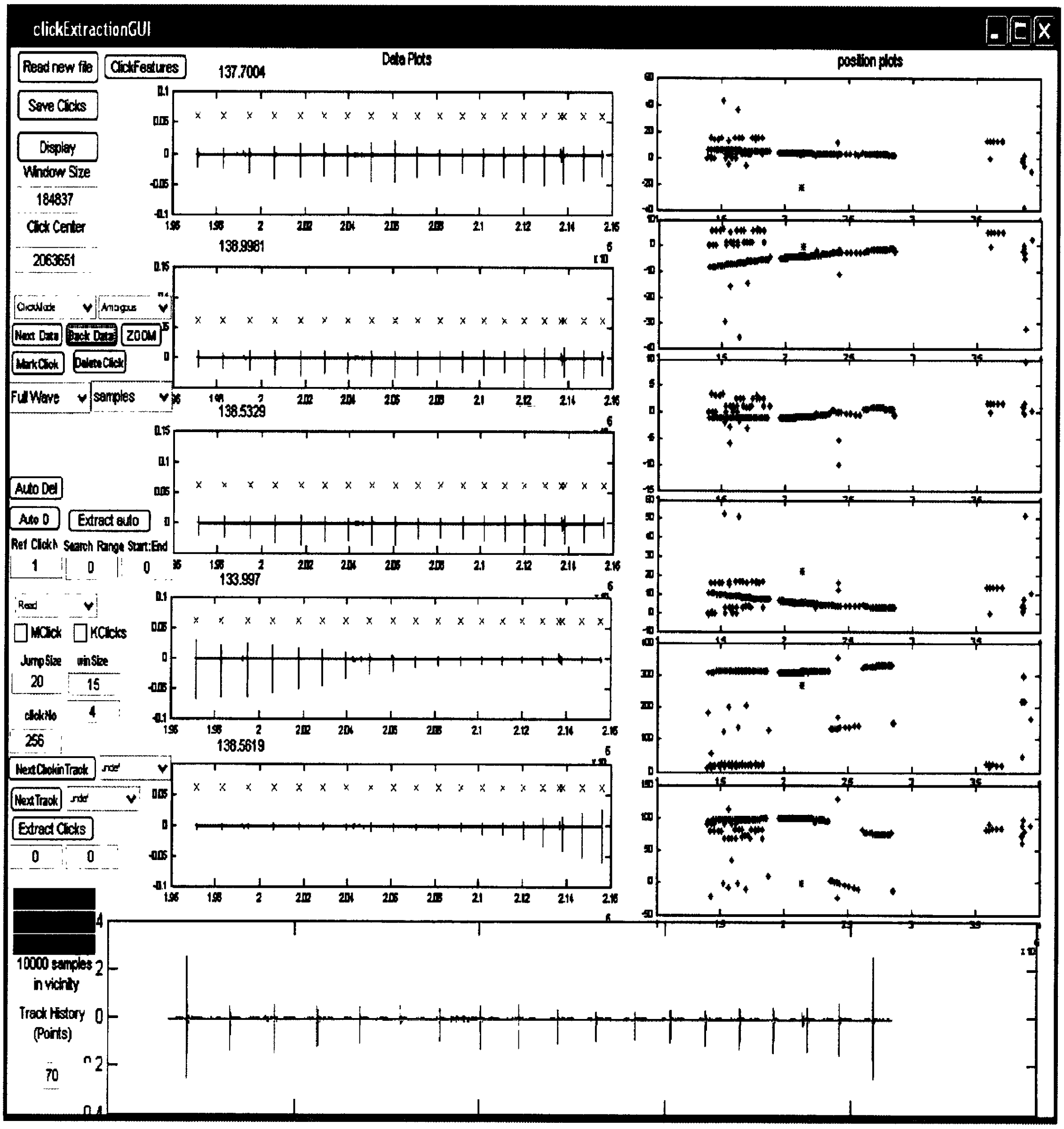

Fig. 6: A snapshot of graphical user interface used in acoustic feature analysis of data of bottlenose dolphins