Development of detection device for dugong calls

RIKA SHIRAKI[1], KOTARO ICHIKAWA[1], TOMIO SHINKE[2], NOBUAKI ARAI[1], TOMONARI AKAMATSU[3], TAKESHI HARA[4] AND KANJANA ADULYANUKOSOL[5]
[1]*Graduate School of Informatics, Kyoto University, 606-8501 Kyoto, Japan*
[2]*System Intech Co., Ltd., R&D Center, 424-8610 Shizuoka, Japan*
[3]*National Research Institute of Fisheries Engineering, 314-0421 Ibaraki, Japan*
[4]*Japan Fisheries Resource Conservation Association, 104-0054 Tokyo, Japan*
[5]*Phuket Marine Biological Center, 83000 Phuket, Thailand*

ABSTRACT
An acoustical approach for research on marine mammals has been a very active research method in recent years. Dugong (*Dugong dugon*) is one of the highly endangered species, which are strictly-marine herbivorous and mainly inhabit coastal areas. In order to detect dugong calls from recorded data, several algorithms have been adapted by researchers in the analyzing process. However, the number of misses in the detection is still non-zero. The sound of snapping shrimp recorded in a wide range (2-300 kHz) is one of the main background noises that makes the detection of dugong calls difficult in warm shallow waters. Impulse elimination was employed in the system to get rid of the snapping shrimp noise. In order to improve the performance of the detection system by increasing the detection rate and decreasing the number of misses, two new algorithms were tested in the experiment. The experimental results for the new algorithms including impulse elimination and the cepstrum method are presented in this paper.

KEYWORDS: dugong, software development, cepstrum

INTRODUCTION

It has been suggested that there is a strong relationship between the vocal exchange of marine mammals and their behavior (Deecke et al., 2002; Janik and Slater, 1998). The functional meanings of their calls now need to be elucidated. Acoustical approaches on marine mammal research has been active (Janik et al., 2000; Anderson and Barclay, 1995; Akamatsu et al., 2001) and become a very fascinating field not only for biologists but also for other scientists. In this study, we have focused on the acoustical data analysis of dugong (*Dugong dugon*) calls and discussed the automatic detection methods for them.

The Dugong is one of the high-degree endangered marine mammal species. They inhabit the coastal areas of the tropics and the subtropics. Several researches have been conducted with the aim of resolving their ecosystem, using passive acoustic methods (Ichikawa et al., 2006). The vocalization of dugongs was collected using stereo-underwater recording systems (AUSOMS-D. Details are described in Ichikawa et al., 2006 and Tsutsumi et al., 2006).

In the process of detecting dugong calls, the researchers listen to the collected data one by one in most cases, and distinguish the objective signals from the background noise. Such listening process obviously consumes an enormous amount of time to organize and analyze the data sets. The purpose of our study is to develop an automatic detection algorithm for dugong calls in order to reduce the listening time for the researchers and improve the efficiency of the data analysis. In this paper, two different algorithms for detecting dugong calls are presented.

1. Cepstrum analysis
2. Impulse elimination method

MATERIALS AND METHODS

Data collection

The data of dugong calls was recorded at Talibong Island, Trang in Thailand from February 28th to March 5th in 2005. AUSOMS-D was employed to collect the sound data. Data analysis was performed by Matlab.

Method

Cepstrum analysis

Cepstrum analysis is one of the most well known analysis methods for human voice recognition. It is also used for seismic echo analysis and other convoluted signal analysis. The ultimate purpose for the use of the cepstrum analysis is to separate the two convoluted signals and to extract the characteristics of the each. In the case of human voice recognition, the characteristics for sound source (vocal cord vibration) and the resonance component (vocal tract configuration) can be extracted, and the envelope of the resonance component is called formant. The steps of cepstrum analysis are as follows,

1. Fast Fourier Transform (FFT)
   FFT was applied to the data in time domain with the window width of 256 point and transformed into the frequency domain signal.
2. Band pass filter (BPF)
   BPF between 1 kHz and 8kHz was applied to the signal since the average frequency of dugong calls was about 4 kHz (Anderson & Barclay, 1995; Ichikawa et al., 2006).
3. Logarithm conversion
   The signal was converted to a logarithm.
4. Invert Fast Fourier Transform (IFFT)
   IFFT was applied to the signal and converted to the quefrency domain signal
5. Liftering and FFT of low quefrency domain
   Liftering window width was set (about 15) and applied FFT to the low quefrency domain signal in order to obtain the envelope of the dugong call.

Impulse elimination

The sound of snapping shrimps recorded in a wide range (2-300 kHz) is one of the main background noises in the data of dugong calls. It makes the detection of dugong calls very difficult in warm shallow waters. Impulse elimination was employed in the system to get rid of the snapping shrimp noise. The steps of the impulse elimination method are as follows.

1. Fast Fourier Transform (FFT)
   FFT was applied to the data in time domain with the window width of 256 and transformed into the frequency domain signal. The shifting was by 128 points.
2. Band pass filter (BPF)
   BPF between 1 kHz and 8 kHz was applied to the frequency domain signal.
3. Power integration
   The sound power was integrated for each window.
4. Threshold and detection of dugong calls
   Set the threshold and declare the signals that are beyond the threshold as dugong calls.

Assessment of detection performance

Detection rate is defined as a ratio of the correct detection counts performed by the system to the human ear based results. False detection rate, on the other hand, is a ratio of the incorrect detection counts performed by the system to the number of candidate dugong calls detected by the system. The receiver operating characteristic curve (ROC curve) can be drawn by calculating both a detection rate and a false detection rate that correspond to different thresholds.

RESULTS

Cepstrum analysis

No peak that indicates the fundamental frequency was observed in the quefrency domain when the dugong calls were applied to the cepstrum analysis (see Fig.1).

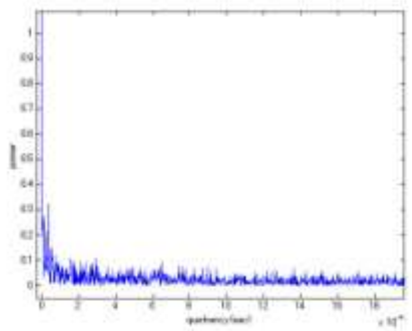

Fig.1 Spectrum of dugong calls in quefrency domain
No obvious peak was observed.

We were able to show an example of the envelope for dugong calls (Fig.2). However, no significant uniqueness was found among those envelopes.

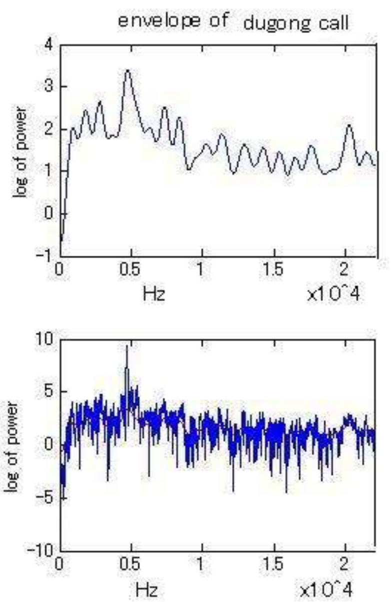

Fig.2 Envelope of log of dugong power spectrum

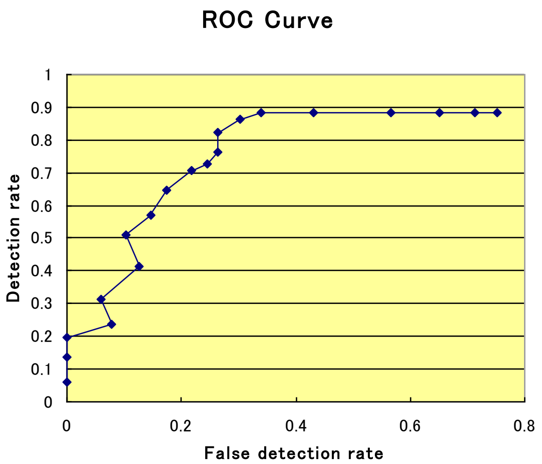

Fig.3 ROC curve obtained from impulse elimination method

Impulse elimination

The detection rate obtained from the elimination method was about 25% and the false detection rate was about 8%. Figure 3 shows the ROC curve evaluating the precision of the impulse elimination method

DISCUSSION

Cepstrum analysis

In the cepstrum analysis, the peak of fundamental frequency (F0) was not distinguishable in the quefrency domain. One of the possible reasons for the lack of the peak is because of the relatively high F0 in dugongs calls. When the F0 is too high, the peak of F0 can be included in the low quefrency domain where the random peaks of noise exist. Another reason for the absence of the F0 peak is the lack of harmonics. Unlike a human voice or other marine mammals' vocalizations, dugongs have only a few harmonics. Therefore, no peak for harmonics can be observed either. The lack of harmonics in dugong calls is a disadvantage in call detection using the cepstrum analysis.

Impulse elimination

How the system should be evaluated depends on the purpose of the system and the requirement of the users. When the role of the detection system is to reduce the listening-time for researchers, the system should be concentrated to reduce the miss rate. In other words, the relatively high false rate can be accepted, in addition to the high detection rate, of course. With our impulse elimination method, the detection rate was 60-70% where the false detection rate was about 20%. On the other hand, when the amount of data is enormous where it is almost impossible for researchers to repeat the listening process, a relatively low false detection rate is required. The detection rate using our method was 20-30% when the false detection rate was less than 10%. Compared to the previous study (Ichikawa el. 2006), the detection rate of our method is still low. Definition of new parameters can be the next step in the further study of dugong detection algorithms. Also, analysis on three dimensional graphs of dugong calls is suggested.

REFERENCES

Akamatsu, T., Wang, D., Wang, K., Wei, Z. (2001). Comparison between visual and passive acoustic detection of finless porpoises in the Yangtze river, China, J. Acoust. *Soc. Am.*, **109**, pp. 1723-1727

Anderson, P., K.. and Barclay, R., M.,R: Acoustic signals of solitary dugongs: Physical characteristics and behavioral correlates, *Journal of Mammalogy*, **76-4**, pp. 1226-1237 (1995)

Deecke V., B., Slater P., J., B, Ford J., K., B.: Selective habituation shapes acoustic predator recognition in harbour seals, Nature, 420-6912, pp. 171-173 (2002).

Ichikawa, K., Tsutsumi, C., Arai, N., Akamatsu, T., Shinke, T., Hara, T., Adulyanukosol, K.: Dugong (Dugong dugon) vocalization patterns recorded by automatic underwater sound monitoring systems, J. Acoust. *Soc. Am.*, **119-6**, pp. 3726-3733 (2006)

Janik, V., M. and Slater, P.,J., B.: Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls, *Anim. Behav*., **56-4**, pp. 829-838 (1998)

Janik, V., M., Van Parijs, S., M., Thompson, P., M.: A two-dimensional acoustic localization system for marine mammals, *Marine mammal science*, **16-2**, pp. 437-447 (2000)