Research Journal of Applied Sciences, Engineering and Technology 4(20): 3911-3917, 2012
ISSN: 2040-7467

Submitted: December 20, 2011 Accepted: April 23, 2012 Published: October 15, 2012

Research on Blind Source Separation of Marine Mammals Signal Processing under Water craft Emitted Noise

Liang Zhang, Longxiang Guo and Jidan Mei
Technology on Underwater Acoustic Laboratory, Harbin Engineering University,
Harbin, Heilongjiang, 150001, China

Abstract: As statistical independence exists between marine mammals sound and watercraft emitted noise and among different organism signals, the study introduces blind source separation into marine mammal signals processing. Blind source separation with single hydrophone is a special underdetermined blind separation problem and blind source separation based on matrix calculating is no longer suitable. The problem can be solved by expanding channels and further study finds that second iteration of blind separation can improve the performance. Algorithm simulation and experimental data analysis show that not only marine mammals signal but also different organism signals can be separated by this method with single hydrophone. It is proved that the correlation coefficient of the separated signal is obviously improved, which lays the foundation for the feature extraction and recognition of marine mammals signal.

Keywords: Blind source separation, marine mammals signal processing, maximum signal noise ratio criterion, second iteration

INTRODUCTION

The many successes of research on marine mammals range from specie diversification, environment analysis, biological evolution and habitable environment (Herman and Clarence, 1998; Castro and Huber, 2011). In retrospect, marine mammals monitoring started in the 1950s and 1960s, monitoring method and data-logger have been developed rapidly over the last several decades (Oswald *et al*., 2004) and a variety of marine mammals sound were recorded and classified in different waters of the world (Zimmer *et al*., 2003). Marine mammal occurrence is currently assessed using visual surveys or passive acoustic monitoring, whereas acoustic monitoring is extensively used as farer survey distance, longer monitoring period and less weather-dependent (Baumgartner and Mussoline, 2011). From 1998 to 2001, 115 h of acoustic recordings were made in the presence of the well-studied St. Lawrence population of blue whales, using a calibrated omnidirectional hydrophone suspended at 50 m depth from a surface isolation buoy (Berchok *et al*., 2006). However, surveys with cabled hydrophones and high-frequency recording devices on board vessels are neither easy to set up nor feasible on many of the vessel research platforms. Songhai Li used a miniature stereo acoustic event data-logger (A-tag) in the survey of the Yangtze finless porpoise and examined the feasibility of using the A-tag on moving vessels for acoustic surveys and fixed platforms for acoustic monitoring of the animal (Li *et al*., 2010). Miller and Dawson (2009) designed a passive sonar array for tracking diving sperm whales in three dimensions from a single small vessel in 2007. The system consisted of four free floating buoys, each with a hydrophone, built-in recorder and one vertical stereo hydrophone array deployed from the boat (Miller and Dawson, 2009).

Traditional acoustic monitoring needs to deploy hydrophones acquiring mammals sound in the waters where marine mammals often occur. The greatest drawback of traditional monitoring is lack of pertinence and efficiency (Baumgartner and Mussoline, 2011; Li *et al*., 2010). Therefore, to obtain real-time marine mammal sounds accurately, research vessels should sail to track marine mammals. Considering the demands of flexibility and maneuverability and the difficulty of large-aperture array deployment, marine mammals sound acquisition and processing are carried out when ships are sailing and the presence of watercraft emitted noise can degrade detection and classification performance, or reduce the effective range of a passive acoustic monitoring system (Gur and Niezrecki, 2009); when the same area existing two or more marine mammals, recorded signal may be mixed by different mammals sound, which brings difficulties to future feature extraction and recognition of marine mammals signal (Stafford *et al*., 2007).

Corresponding Author: Liang Zhang, Technology on Underwater Acoustic Laboratory, Harbin Engineering University, Harbin, Heilongjiang, 150001, China

As statistical independence exists between marine mammals sound and watercraft emitted noise and among different organism signals, Blind Source Separation (BSS) is suitable for marine mammals signal processing. Despite the extensive literature on speech and communications source separation, only a few studies implement source separation in the context of underwater acoustics (Gur and Niezrecki, 2009). Based on the present research, a method is proposed to separate marine mammals signal with single hydrophone by expanding channels. Algorithm simulation and experimental data analysis show that not only marine mammals signal but also different organism signals can be separated by this method with single hydrophone. It is proved that the correlation coefficient of the separated signal is obviously improved, which lays the foundation for the feature extraction and recognition of marine mammals signal.

METHODOLOGY

Blind source separation algorithms: BSS is a class of adaptive signal processing that serve for retrieving the original signals emitted from multiple point sources from multi-channel mixtures. Referred to as blind, both the source signals and the mixing channel are assumed to be unknown. The task of BSS is to recover original sources from their linear instantaneous mixtures without resorting to any prior knowledge except for the statistical independence of sources. The linear BSS problem is expressed as:

$$X(n) = Hs(n) \tag{1}$$

where, x(n) is the N-dimensional observation vector, H is the N×N instantaneous linear mixed matrix and s(n) is the N-dimensional vector of sources that are assumed to be mutually independent.

For source separation, a demixing matrix w is to be found by prior knowledge of probability distribution about observation signal x(n) and source signal s(n), such that $w^H H = I$. This gives:

$$y(n) = Wx(n) = Whs(n) \tag{2}$$

where, y(n) is the estimate of s(n), namely estimation signal or separation signal.

Blind source separation algorithms based on the maximum signal-to-noise ratio: BSS algorithms can be separated into several different groups based on the implementation of the statistical independence criterion and we select the algorithms based on Maximum Signal-to-Noise Ratio in this study. SNR function is established as follows:

$$SNR = 10\log\frac{s.s^T}{e.e^T} = 10\log\frac{s.s^T}{(s-y).(s-y)^T} \tag{3}$$

where, s is the sources, y is the estimate of s and the error term e defined as e = s−y.

As the source s is unknown and y(n) includes the contributions of noises, moving average $\tilde{y}$ of y(n) instead of s and then (3) turns into:

$$SNR = 10\log\frac{s.s^T}{e.e^T} = 10\log\frac{\tilde{y}.\tilde{y}^T}{(\tilde{y}-y).(\tilde{y}-y)^T} \tag{4}$$

where, $\tilde{y}_i(n) = \frac{1}{p}\sum_{j=0}^{p} y_i(n-j), i = 0, 1,2, \ldots, p-1$, sliding average length p can be chosen according to the signal noise characteristic (p can be positive integer less than 100). To simplify the calculation, $\tilde{y}$ instead of y in (4). Thus the objective function of maximum signal-to-noise ratio is obtained as:

$$F(y) = SNR = 10\log\frac{y.y^T}{(\tilde{y}-y).(\tilde{y}-y)^T} \tag{5}$$

where, y = Wx, $\tilde{y} = W\tilde{x}$, w is the demixing matrix. And the mixed signal $\tilde{x}$ after sliding average can be expressed as:

$$\tilde{x}_i(n) = \frac{1}{p}\sum_{j=0}^{p} y_i(n-j), x \in i = 0, 1, 2, \ldots, p-1 \tag{6}$$

Equation (5) can be written as:

$$\begin{aligned} F(W, x) &= 10\log\frac{y.y^T}{(\tilde{y}-y).(\tilde{y}-y)^T} \\ &= 10\log\frac{Wxx^TW^T}{W(\tilde{x}-x)(\tilde{x}-x)^TW^T} \\ &= 10\log\frac{WCW^T}{W\tilde{C}W^T} = 10\log\frac{V}{U} \end{aligned} \tag{7}$$

where, $C = xx^T$, $\tilde{C} = (\tilde{x}-x)(\tilde{x}-x)^T$, $V = WCW^T$, $U = W\tilde{C}W^T$.

Both sides of (7) to calculate the gradient using the demixing matrix w:

$$\frac{\partial F}{\partial W} = \frac{2W}{V}C - \frac{2W}{U}\tilde{C} \tag{8}$$

Since the extreme value point of target function F(W, x) is zero point of (8), we can get:

$$WC = \frac{V}{U}W\tilde{C} \tag{9}$$

The solution $\hat{w}$ of (9) is the eigenvector of the matrix $\tilde{C}.C^{-1}$. The vector of separate source signal is $y = \hat{W}x$, where each line of y presents a separate signal and y_i is the enhanced estimate of the source.

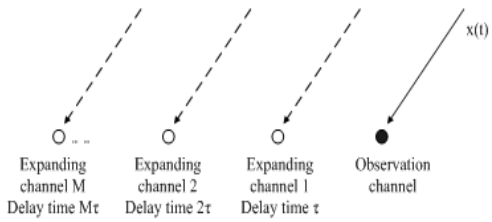


Fig. 1: The schematic of BSS based expanding channels

Evaluation of performance: In order to evaluate the performance of the BSS algorithms, the correlation coefficient ζ_{ij} of separate signal and source signal is defined as performance criteria, that is:

$$\zeta_{ij} = \varsigma(y_i, s_j) = \left|\sum_{t=1}^{M} y_i(t) s_j(t)\right| \Big/ \sqrt{\sum_{t=1}^{M} y_i^2(t) \sum_{t=1}^{M} s_j^2(t)} \tag{10}$$

when $y_i = cs_j$, c, is constant, $\zeta_{ij} = 1$, in other words, the estimates have amplitude differences with sources; When $\zeta_{ij} = 0$, y_i and s_j are mutual independence. That is to say, each line and each column of the matrix constituted by correlation coefficient has one and only one element approximation close to 1, meanwhile other elements are similar to 0 and the performance can thought to be ideal.

BSS based expanding channels: As the noise levels increase relative to the source strength, a signal enhancement procedure becomes necessary. The application of BSS to marine mammals sound, especially the presence of dominant noise sources, is a challenging task because of the active sources contrasting with the relative paucity of hydrophones (Gur and Niezrecki, 2009). Usually marine mammals acoustic monitoring uses single hydrophone to collect mammal sounds. BSS with single hydrophone is a special underdetermined blind separation problem and BSS based on matrix calculating is no longer suitable. The study proposes a method to solve the problem by expanding channels.

Figure 1 depicts the schematic of BSS based expanding channels. The black point is the actual observation channel with hydrophone; other points are expanding channels which not exists. In order to have a better comprehension, it can be considered as a uniform linear array and the actual hydrophone is at the far right. Assuming the number of sources is N, the actual hydrophone output is:

$$x(t) = \sum_{i=1}^{n} a_i s_i(t) \tag{11}$$

It is assumed that the time delay is defined as τ between every two channel, which beyond correlation radius of noise and within the range of signal. The performance of the algorithm is sensitive to time delay and the value of time delay is chosen to get the best performance in this study. In practical engineering, we should pay close attention to time delay as it severely affects the performance. Therefore the received signals of expanding channels can be expressed as:

$$x_j(t) = x(t+j\tau)\ j\epsilon[1, M] \tag{12}$$

And the matrix of received signals defined as:

$$x(t) = \begin{bmatrix} x(t) \\ x(t+\tau) \\ \vdots \\ x(t+M\tau) \end{bmatrix} \tag{13}$$

Typically it is required that the number of sources is less or equal to the number of observation channels, $M \geq N-1$. In that case, the linear mixture is invertible and the source signals can be recovered. After the above transformation, blind source separation based on matrix calculating is suitable for (13).

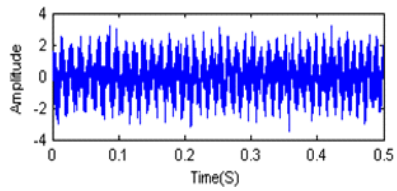
SIMULATIONS

This section investigates the performance of the proposed method and the validity of the method is also assessed. To evaluate the performance of the algorithm under ideal conditions, source signals are generated. Sources generation:

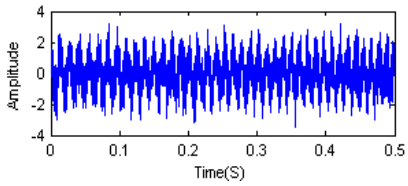
- Linear Frequency Modulation signal (LFM), the initial carrier frequency is 50 Hz and frequency modulation slope of LFM is 40 Hz/s
- Band-limited Gaussian white noise, band width is 10 Hz~5 kHz

As the number of sources is less or equal to the number of observation channels, the number of expanding channels is M = 1. And time delay between actual channel and expanding channel is τ = 0.6 ms, which beyond correlation radius of band-limited Gaussian white noise and within the range of LFM. The Signal-to-Noise Ratio (SNR) is defined as the power rate of LFM and limited Gaussian white noise. In the context of signal enhancement, separation of LFM and band-limited Gaussian white noise demands an improvement in the input SNR.

Figure 2 depicts the waveform of the mixture of the sources when SNR = 0 dB. As the interference of the band-limited Gaussian white noise, we can just recognize the outline of LFM. Figure 2a shows the acquisition signal of hydrophone, meanwhile, Fig. 2b presents the acquisition signal of expanding channel which lags 0.6 ms referred to Fig. 2a. In that case, the separation signals are shown in Fig. 3. Figure 3a depicts the estimate of band-l

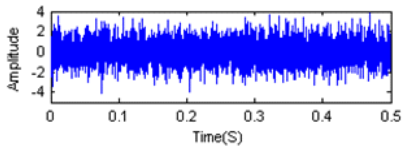


(a)

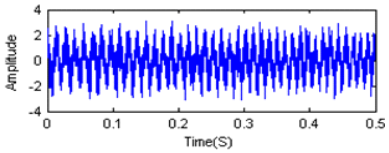


(b)

Fig. 2: The time domain plot of a linear mixture of the two sources, (a) The waveform of the acquisition signal of the hydrophone, (b) The waveform of the acquisition signal of the expanding channel



(a)



(b)

Fig. 3: The estiates of the sources (SNR = 0 dB), (a) Band-limited gaussian white noises, (b) LFM

imited Gaussian white noise and Fig. 3b depicts the estimate of LFM. From Fig. 3 we can see that LFM is separated from limited Gaussian white noises and the correlation coefficient of LMF and the estimate rises from 0.7098 to 0.8082 after the BSS.

Further study founds that Second-Iteration BSS (SIBSS) has a better performance than First-Iteration BSS (FIBSS). SIBSS considers the output of FIBSS as the source. Figure 4 shows the performances of FIBSS and SIBSS under different SNR varying from -10 to 10 dB. It is clear from Fig. 4, the correlation coefficient continually increases with the development of SNR and the performance index of SIBSS is obviously higher than that of FIBSS. However, when SNR at a high level (SNR>8 dB) the difference of performance index is no longer apparent but SIBSS still maintains the advantage.

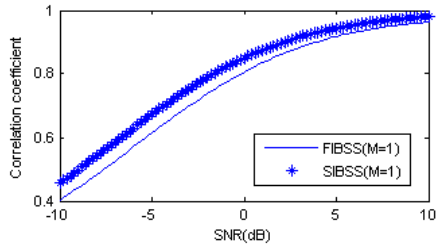
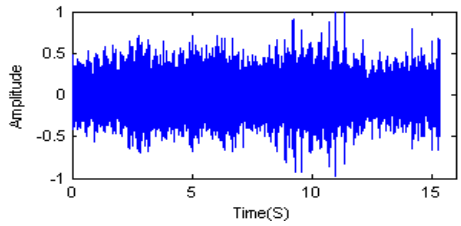
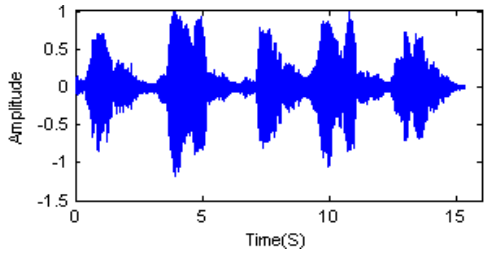


Fig. 4: The performances of FIBSS and SIBSS under different SNR



(a)

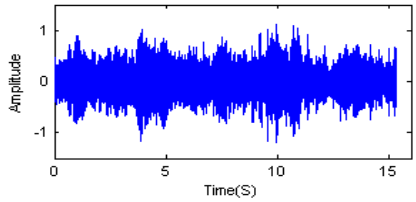


(b)

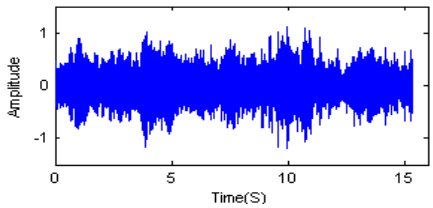
Fig. 5: The time domain plot of source signals, (a) Watercraft emitted noise, (b) The dolphin vocalizations

Experimental analysis: On the basis of theoretical derivation and simulations, the handling ability of algorithm to experimental data is studied in this section. To simplify the complexity of the algorithm, the sources are the mixture of two original signals. The experiment is divided into two groups and the first group adopts the linear mixture of watercraft emitted noise and dolphin vocalizations as source signals, while the other group chooses different dolphins vocalizations as source signals. Watercraft emitted noise is a kind of radiation noise that collected locally as watercrafts navigating; the dolphin vocalizations are obtained under static drift. A typical dolphin vocalization lasts between 1 and 2s and may have several harmonics in the frequency band of 5-20 kHz. As the number of each experiment source signals is 2, the number of expand channels is chosen M = 1.

Experiment 1: BSS based expanding channels of the linear mixture of watercraft emitted noise and dolphin vocalizations. Figure 5 depicts the waveform of source signals. Figure 5a shows the watercraft emitted noise

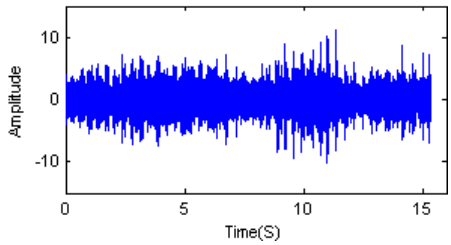


(a)

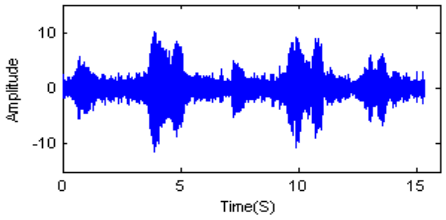


(b)

Fig. 6: The time domain plot of a linear mixture of the two sources, (a) The waveform of the acquisition signal of the hydrophone, (b) The waveform of the acquisition signal of the expanding channel



(a)



(b)

Fig. 7: The estimates of the sources (SNR = 0 dB), (a) The dolphin vocalizations, (b) Watercraft emitted noise

recorded locally as watercrafts navigating and Fig. 5b presents the time domain plot of a dolphin vocalization as the ambient noise is very weak and there are no other dolphins exist to get the pure dolphin sounds. The linear mixture of the sources is shown in Fig. 6 at SNR = 0 dB, in which we couldn't recognize the outline of the sources. Figure 6a shows the acquisition signal of hydrophone, meanwhile, Fig. 6b presents the acquisition signal of expanding channel which lags 0.8 ms referred to Fig. 6a.

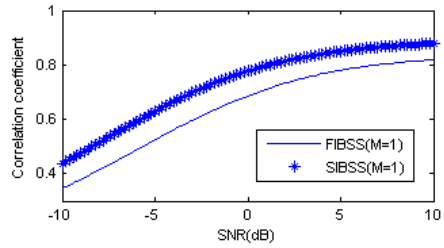
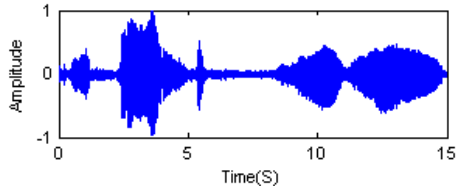
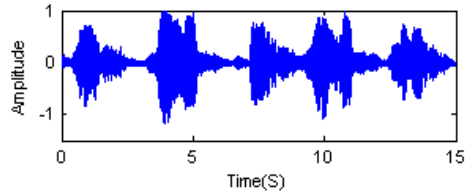


Fig. 8: The performances of FIBSS and SIBSS under different SNR



(a)

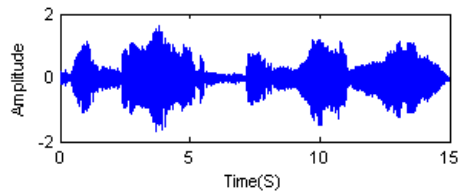


(b)

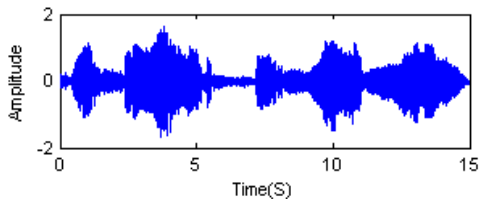
Fig. 9: The time domain plot of source signal, (a) Whiles, (b) Clicks

Figure 7 depicts the estimates of the sources of SIBSS and Fig. 7a shows the estimate of dolphin vocalization, meanwhile Fig. 7b is the estimate of watercraft emitted noises. Comparing the time domain plot of Fig. 5 and 7, the dolphin signal has been separated from the watercraft emitted noise. Existing a problem that water emitted noise is much different from its estimate. The source signal of SIBSS is the estimate of the dolphin vocalization which from FIBSS. The dolphin vocalization is separated from the watercraft emitted noises in some degree after FIBSS when SNR = 0 dB and the dolphin vocalization is dominant in the estimate. Therefore, SIBSS means to separate the signal from the pure sources and it is not surprising that both of the estimates are similar with the dolphin vocalization. It is an available method of changing the proportion of watercraft emitted noise in the mixture signals to get the estimate of watercraft emitted noise. The problem demonstrates the validity of the algorithm based expanding channels.

Figure 8 shows the comparison of performances of FIBSS and SIBSS under different SNR varying from -10 to 10 dB. It can be seen that the performances of FIBSS and SIBSS have been improved with the rising SNR and

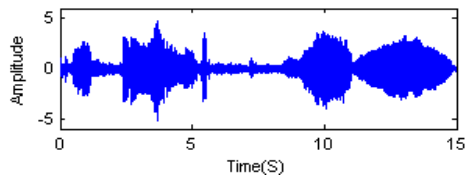


(a)

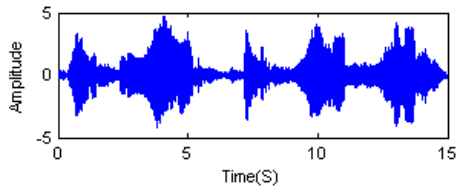


(b)

Fig. 10: The time domain plot of a linear mixture of the two sources, (a) The waveform of the acquisition signal of the hydrophone, (b) The waveform of the acquisition signal of the expanding channel



(a)



(b)

Fig. 11: The estimates of the sources (SNR = 0 dB), (a) Whiles, (b) Clicks

the correlation coefficient of SIBSS is higher approximate to 0.1 than that of FIBSS. The above analysis shows that SIBSS will lose estimates to one source signal, however, can significantly improve the estimated performance to the interesting signal.

Experiment 2: BSS based expanding channels of the linear mixture of two dolphin vocalizations. Figure 9 depicts the time domain plot of the sources. Dolphin vocalizations can be divided into two classes: whistles and clicks. Whiles are widely spread across species, playing crucial social role and thus should be termed "calls" and clicks are used for echolocation. Figure 9a

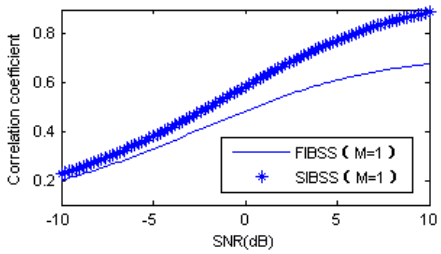


Fig 12: The performances of FIBSS and SIBSS under different SNR

shows the waveform of whiles and Fig. 9b presents the clicks. The linear mixture of the sources is shown in Fig. 10 at the SNR = 0 dB,τ = 0.9ms.

Figure 11 depicts the estimates of the sources of SIBSS. Figure 11a shows the estimate of whiles, meanwhile Fig. 11b is the estimate of clicks. Comparing the time domain plot of Fig. 9 and 11, the sources are separated effectively. Figure 12 shows the comparison of performances of FIBSS and SIBSS under different SNR varying from -10 to 10 dB. It can be seen that the performances of FIBSS and SIBSS have been improved with the rising SNR and the correlation coefficient of SIBSS is higher than that of FIBSS especially SNR is high.

CONCLUSION

As statistical independence exists between marine mammals sound and watercraft emitted noise and among different organism signals, the study introduces BSS into marine mammals signal processing.

BSS with single hydrophone is a special underdetermined blind separation problem and BSS based on matrix calculating is no longer suitable. The study presents the BSS algorithm based expanding channels can overcome the problem. Simulations and experimental analysis indicate the algorithm is an effective method to the special underdetermined blind separation problem and sources can be separated effectually, particularly at high input SNR values. And further study finds that second iteration of blind separation can evidently improve the performance.

The BSS algorithm based expanding channels has two advantages in marine mammals signal processing. First, marine mammals sound separation is realized using just one hydrophone which results in considerable simplification of the equipments and lower research cost. Secondly, not only marine mammals signal and watercraft emitted noise but also different organism signals can be separated by this method with single hydrophone, which lays the foundation for the feature extraction and recognition of marine mammals signal.

ACKNOWLEDGMENT

The authors would like to gratefully acknowledge the financial support of the National High Technology Research and Development Program 863 (2006AA09Z234), the Science and Technology on Underwater Acoustic Laboratory (9140C200505080C20).

REFERENCES

Baumgartner, M.F. and S.E. Mussoline, 2011. A generalized baleen whale call detection and classification system. J. Acoust. Soc. Am., 129: 2889-2902.

Berchok, C.L., D.L. Bradley and T.B. Gabrielson, 2006. St. Lawrence blue whale vocalizations revisited: Characterization of calls detected from 1998 to 2001. J. Acoust. Soc. Am., 120: 2340-2354.

Castro, P. and M.E. Huber, 2011. Marine Biology. 6th Edn., Peking University Press, Beijing, pp: 438-446.

Gur, M.B. and C. Niezrecki, 2009. A source separation approach to enhancing marine mammal vocalizations. J. Acoust. Soc. Am., 26: 3062-3070.

Herman, M. and S.C. Clarence, 1998. Fundamentals of Acoustical Oceanography. Academic Press, San Diego, pp: 348-362.

Li, S.H., T. Akamatsu, L.J. Dong, K.X. Wang, D. Wang and S. Kimura, 2010. Widespread passive acoustic detection of Yangtze finless porpoise using miniature stereo acoustic data-loggers: A review. J. Acoust. Soc. Am., 128: 1476-1482.

Miller, B. and S. Dawson, 2009. A large-aperture low-cost hydrophone array for tracking whales from small boats. J. Acoust. Soc. Am., 126: 2248-2256.

Oswald, J.N., R. Shannon and B. Jay, 2004. The effect of recording and analysis bandwidth on acoustic identification of delphinids species. J. Acoust. Soc. Am., 116: 3178-3185.

Stafford, K.M., D.K. Mellinger, S.E. Moore and C.G. Fox, 2007. Seasonal variability and detection range modeling of baleen whale calls in the Gulf of Alaska, 1999-2002. J. Acoust. Soc. Am., 122: 3378-3390.

Zimmer, W.M.X., M.P. Johnson, A.D.Amico and P.L. Tyack, 2003. Combining data from a multisensor tag and passive sonar to determine the diving behavior of a sperm whale (*Physeter macrocephalus*). IEEE J. Oceanic Eng., 28: 13-28.