# A Compact Real-Time Acoustic Bandwidth Compression System for Real-Time Monitoring of Ultrasound

Teong Beng Koay[1], John Robert Potter[1], Mandar Chitre[1],
Sheldon Ruiz[1], Eric Delory[2]

[1]Acoustic Research Laboratory, Tropical Marine Science Institute,
National University of Singapore, 12a Kent Ridge Road, Singapore 119223.
[2]Laboratory of Applied Bioacoustics, Universitat Politecnica de Catalunya, Spain
koay@arl.nus.edu.sg

***Abstract* - Many animals and systems radiate ultrasound that contains valuable information, from bats to high-voltage power lines. We set out to develop a real-time bandwidth compressor that can convey the prominent features of ultrasound in the human hearing band of 50 Hz to 16 kHz that requires no assumption of where in the ultrasonic frequency band of interest or when in the time domain the information is encoded.**

**A primary application is in dolphin communication, which is believed to be both sophisticated and ultrasonic. Real-time studies of their acoustic communication patterns together with their behavior with the added capability to be able to react and respond in a timely manner to interact with them could rapidly generate important findings, greatly improving the efficiency of dolphin-human interactions. This is not possible without a real-time interface between ultrasound and human hearing.**

**As is well known, there is no pictorial representation that readily conveys the richness of a sound. We are therefore driven to find an efficient acoustic interface, translating ultrasound into audible sounds. This is easy to do in post-processing (simply play back at reduced speed), but continuous streaming real-time processing presents a challenge. The total information-carrying capacity of a signal can be represented by the time-bandwidth product. If the time is constrained to be the same and the bandwidth must be reduced, some information must be discarded. Choosing how and where to do this is the key to a successful algorithm.**

**In this paper we present an algorithm that compresses ultrasound signals into the audio band of human hearing while maintaining the overall signatures and structures of the signal, regardless of the signal type. This algorithm can be demonstrated to be optimal under the applied constraints. This is followed by the design of a prototype system that provides real-time bandwidth compression and a preliminary test result of the system capability.**

**The algorithm has a time-domain implementation that makes it possible to downshift signals sampled at up to 1MSa/s to audio range using a DSP. The system is autonomous and compact so that it can be carried by operators, including divers, allowing them to swim among dolphins while listening to their communications. The system is demonstrated using high frequency acoustic signals from a bottlenose dolphin.**

## I. INTRODUCTION

Animals generally post gesture cues in their communication either vocally or physically [1] [2], hence being able to simultaneously observe their vocal and physical behaviours would increase our understanding to their communication. Similar studies have been done with many primates [3]. As many animals have sound production and perceiving capabilities in ultrasound that potentially useful for communication [4] [5], studying vocalisation in the ultrasound range is important. Nevertheless, such a study is difficult to conduct for animals that communicate at ultrasound range unless we are able to effectively perceive ultrasound. This paper describes an algorithm and device designed for this purpose.

Playback-and-study experiments have been carried out to correlate acoustic cue with the animal behaviours in order to improve our understanding to their communications and behaviours [6] [7]. Although this technique is widely employed in many animal communication and behavioural studies [8], it is generally lack of intuitive interactions between researchers and the animals. We expect that being able to ‘listen’ to the animals’ acoustic cues in real-time and interact with them would generate important findings. This translates to the requirement of a system that would transform the signature (frequency, amplitude and time line) of ultrasound to human’s audible range in real time.

Many techniques have been used in order to detect and transform ultrasound into human audible band [9].The functionality of these techniques varies between a real-time detector at one end of the functionality spectrum, to time dilation system that loses real time capability at the other end. The real-time detectors (such as ultrasound alarm, superheterodyne detector, frequency division detector) generally suffer major information losses whereas the time dilation but time dilation systems retain full information of the captured signal.

Basic superheterodyne detector needs prior information about the ultrasounds in order to tune a frequency band (normally same bandwidth of human audio band) around it and downshift it to audio range. The downside is any signal outside the band is lost.

More advanced heterodyne techniques may produce simultaneous tuning capability of multiple bands but blocks of spectral frequency information are discarded. Basic frequency division (zero-crossing division) on the other hand, able to map a range of ultrasound frequency into audio band by dividing the frequency but it loses most of the amplitude information. Although new techniques have provided some capabilities of preserving the amplitude, the system still suffers from the potential of creating wrong downshifted frequency of moving source and missing multiple frequencies.

If the amount of information within a signal can be represented by its time-bandwidth product [10], then when the spectrum frequency is down-shifted, time has to be expended to display the full information. If time dilation is not an option (such as in real-time applications), then some information has to be discarded either in the time or frequency domain. Our Acoustic Bandwidth Compression (ABC) approach is also bounded to the same physical limitation; some information has to be discarded in order to ensure real time mapping. What makes it different from other methods is that it assumes nothing about the ultrasound signal but still preserves the overall 'signature' of the ultrasound, both in time and in frequency.

## II. APPROACH

The ABC algorithm (patent pending) divides the ultrasound time series into segments, compresses their bandwidth and concatenates the segments to resemble a new but representative time series in audio band. Two approaches that eventually lead to a convergence in implementation were designed, one is to compress in the frequency domain while the other is to compress in the time domain. The time domain algorithm has the advantage of reduced computational load. As a result, it reduces signal processing power requirement; hence making it suitable for a real-time, battery-operated system.

### *A. The frequency domain algorithm*

In this approach, the compression is performed in the frequency domain, where we transform the segments of time series into the frequency domain, compress the frequency band, transform it back to time domain, and play back with the new (reduced) sampling rate. As the signal was divided into short discrete sections Discrete Windowed Fourier Transforms (DWFT) and inverse-DWFT are used.

In order to compress the frequency band, frequency estimates are grouped and each group is replaced with a value that is representative both in the amplitude and phase information of the particular group. The number of frequency estimates in each group will be determined by the desired frequency compression ratio. Since we assume nothing about the signal, all the frequency estimates are treated with equal priority. Therefore the frequency estimates are grouped evenly across the frequency band. The result is that the high-resolution frequency structures are removed and the gross signature of the signal is retained.

This is achieved by convolving the frequency estimates with a rectangular window followed by reducing the sampling frequency accordingly. The convolution process is effectively performing low pass filter to the frequency estimates, where length of the windows is relative to the compression ratio desired.

### *B. The time domain algorithm*

The convolution process in frequency domain with rectangular window is equivalent to multiplying the time series with a sinc function in time domain. This results in each section with spurious oscillations at both ends, which needs to be minimised.

Recall that the process is a low pass filter process in frequency domain followed by spectral sub-sampling. As an ideal low-pass filtering in time domain should produce a sharp, vertical cut-off in frequency domain, the ideal frequency domain low pass filter in our approach should yield a rectangular multiplication window in the time domain starting from the beginning of each section, as shown in the first block of Fig. 1.

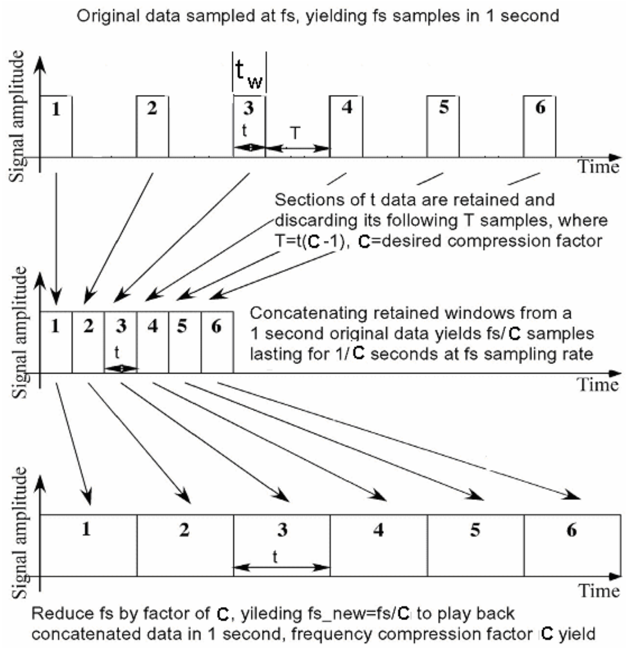


Fig. 1: Time domain algorithm schematic

Fig. 1 shows the summary schematic of the time domain approach. The algorithm simply retains a section of the samples; discards a section of subsequent samples while the sampling rate is reduced accordingly. This is repeated through out the entire time series. The ratio between the retained section and discard sections is related to the compression ratio.

*C. Concatenation*

If left with coarse transitions, the concatenation process nearly always produces spurious noise due to discontinuity between the edges of subsequent sections that are joined together.

This unwanted defect is significantly reduced in two ways. The first method is to allow some flexibility when defining the retention sections in the time series to minimize the level difference. The other method is to apply a smoothing window to each segment and introduce overlaps when concatenating them. This has resulted in significant reduction of the spurious noise.

*D. Performance enhancement by thresholding*

When the input signal contains sparse ultrasound pulses, the sample-compress-concatenate process with regular interval might miss them. In order to avoid this, we search for regions within the time series that contain energy level above a threshold and define the position of the retention window there. Hence, we make sure that regions with high energy content are compressed. This will not only minimise the chances of missing short pulses, but also reduce the computation effort when the input signal is relatively 'quiet'.

## III. PARAMETERS OF ABC

The ABC algorithm does not require detailed prior knowledge of the ultrasound when compressing the bandwidth. Nevertheless, the compression parameters need to be tuned in order for the algorithm to work optimally and safely. This is due to the limitation of human hearing physiology. For example, the compression parameters should produce audible signals with energy levels that are detectable by human ear without damaging it.

These parameters includes the compression factor (C), gain factor (G), the length of the retention section ($t_k$), and the length of the smoothing window ($t_c$). The following discussions describe the tuning of the parameters.

The compression factor will determine the ultrasound bandwidth that the user wants to map into human hearing range. As the compression factor will determine how much high-resolution frequency structure is to be discarded, we would keep the compression factor to minimum. Hence we would ensure that the compression maps the highest frequency of interest into the higher end of the user's audible frequency range. Thus,

$$C = \frac{f_{hi}}{f_{\max}} \tag{1.0}$$

where,

$f_{hi}$ is the upper end of the ultrasound bandwidth

$f_{max}$ is the maximum frequency the user can hear

When determining the length of the retention window and smoothing window, we ensure that the length of the output audio block is at least equal to the human hearing integration time, $\tau$ (see (2.0)). This is to ensure that the user could integrate the acoustic power over the period and gain enough energy to sense the sound even the acoustic level is smaller than detection threshold. Nevertheless, we keep the retention window length, $t_k$, as small as possible to avoid loss of ultrasound transient patterns. The length of the smoothing window, $t_c$, is kept smaller or equal to $t_k$, in order to make sure that the main energy content of compressed signal is from the retention window.

$$(t_k + t_c)C \geq \tau \tag{2.0}$$

The gain setting needs to be large enough so that the energy content of the compressed signal produced from the weakest detectable (by hardware) ultrasound is at least the minimum energy detectable by the human ear over the integration time.

On the other hand, there could be occasions when the input pulse width is smaller than hearing integration time even after compression. Hence we also need to ensure that the smallest acoustic signal of interest is amplified to our minimum hearing threshold. Therefore the gain can be written as,

$$\mathrm{G} = \max\left(\frac{\mathrm{P_a}}{\mathrm{P_{lo}}}, \frac{\mathrm{E_0}}{\mathrm{CP_{lo}\Delta t}}\right) \tag{3.0}$$

where,

$P_a$ is the minimum audible sound pressure for human ear

$P_{lo}$ is the minimum audio level that the hardware can detect

$E_0$ is the minimum audible energy in hearing integration time

$\Delta t$ is the smallest ultrasound pulse width one wish to detect

The above description was demonstrated to be the method of tuning the algorithm for optimum performance. The parameter tuning is mainly related to human hearing physiology, other from that, the other information needed is the ultrasound frequency band of interest.

## IV. PERFORMANCES VARIATIONS DUE TO PARAMETER CHANGES

An ultrasound sweep was used to evaluate the algorithm performance at different parameter settings. The performance statistics is the 2-dimensional zero-lag correlation of the spectrograms of the original and compressed signals, which is normalised by the energy content of both signal. Equation (4.0) describes the process,

$$R_{AB} = \left( \frac{\sum_m \sum_n \left(A_{mn} - \overline{A}\right)\left(B_{mn} - \overline{B}\right)}{\sqrt{\left(\sum_m \sum_n \left(A_{mn} - \overline{A}\right)^2\right)\left(\sum_m \sum_n \left(B_{mn} - \overline{B}\right)^2\right)}} \right) \qquad (4.0)$$

where,

$R_{AB}$ is the correlation coefficient

$A$, $B$ are spectrogram matrix (absolute value) of original signal and compressed signals respectively

$m$, $n$ are the column and row size of the matrix

$\overline{A}, \overline{B}$ are the mean of the spectrogram matrix $A$ and $B$ respectively

A single 12-second frequency sweep from 80 kHz to 120 kHz at a signal to noise ratio of 55dB is used. The compression parameters are maintained at constant values of retention window length of 10ms, smoothing window length of 0.5ms, and compression factor of 10 times except the particular parameter under investigation (one at a time). The gain is kept constant throughout the tests.

The tests provide an idea of how the system behaves at different values of the parameters but does not signify the absolute performance of the algorithm. This is because the human brain can be surprisingly efficient at picking up the audio patterns in highly noisy environment. For example, the authors can still audibly recognise the presence of the acoustic pattern in compressed signals although the value of normalised cross-correlation is less than 0.2.

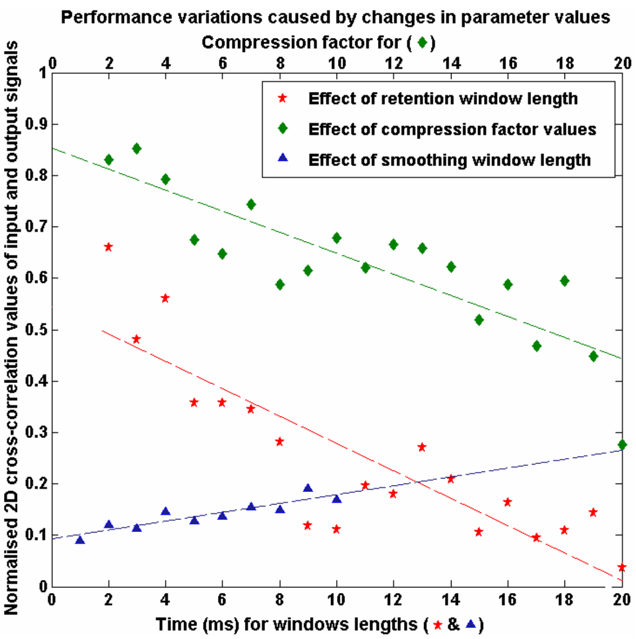


Fig. 2 Variations of algorithm performance affected by changes in parameters values

Fig. 2 shows the performance variations of the system. The dotted lines are the respective linear fits to the curves. It is observed that reducing both the compression factor and the length of retention window would yield better performance than increasing it. This is understandable because larger values in both parameters would result in larger data section to be discarded each time. Thus introducing more mismatch between the original and compressed signals. Note that the fit to the retention window length tests ends at around 1ms. This is because the algorithm needs a minimum number of samples in the retention window to perform the compression and applying smoothing windows.

On the other hand, the larger the length of smoothing window, the more overlaps of time series will be included and lesser discontinuity exist in the compressed signal. This reduces the overall mismatch between original and compressed signal.

Therefore it is seems desirable to keep the retention window and compression factor as small as possible but use larger smoothing window.

Fig. 3 shows a compression example of a test signal. The test signal is 12 seconds long with sampling frequency of 500 kHz. It contains frequency sweeps with harmonics. The audio band of the original signal is filtered with a 6 pole elliptical high pass filter with cut off frequency at 25 kHz. The bandwidth of the test signal is 225 kHz, spanning from 25 to 250 kHz.

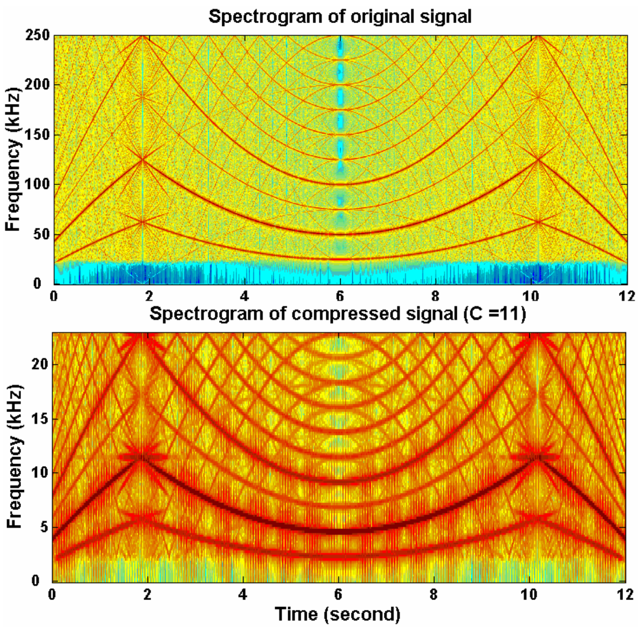


Fig. 3: Spectrogram of a test signal compressed with 1:11 compression ratio

It is observed that the entire time-frequency pattern of the original signal is compressed and fitted to the 0 to 22.7 kHz band. It is to be noted that the shape of the original spectrogram (with the presents of simultaneous-multiple frequencies at any point of time) is preserved. The output spectrogram also shows spectral leakage during concatenation of the compressed blocks, which can be reduced by choosing larger smoothing window.

## V. RESULTS WITH DIFFERENT INPUT SIGNALS

This section presents some test results of the acoustic bandwidth compression using different ultrasonic bio-acoustic signals.

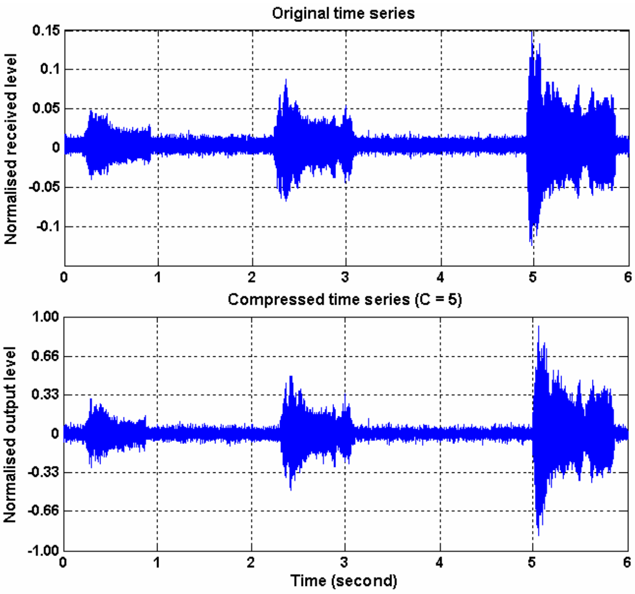


Fig. 4 Comparison between signal shapes of original and compressed pilot whale whistle

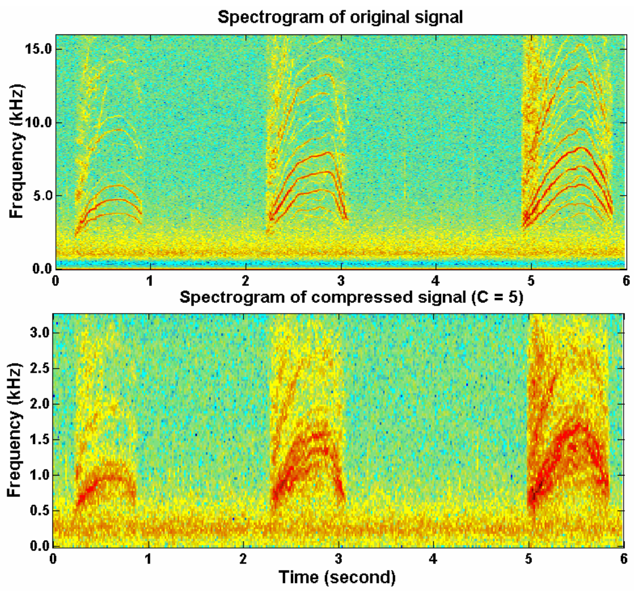


Fig. 5 Comparison between spectrograms of original and compressed pilot whale whistle

The first test signal used is a pilot whale whistle[1] sampled at 32ksps. The specificity of this signal is that it is relatively long compared to the segment of the time series (refer to Fig. 1, it is t+T). The signal is compressed 5 times. The time series of the original and compressed signals are given in Fig. 4. At first sight, the shape of the amplitude envelope follows the original signal closely. Fig. 5 shows the spectrogram of both original and compressed signals. Again, the shape of the overall signal signature is maintained over time and frequency despite the loss of frequency resolution.

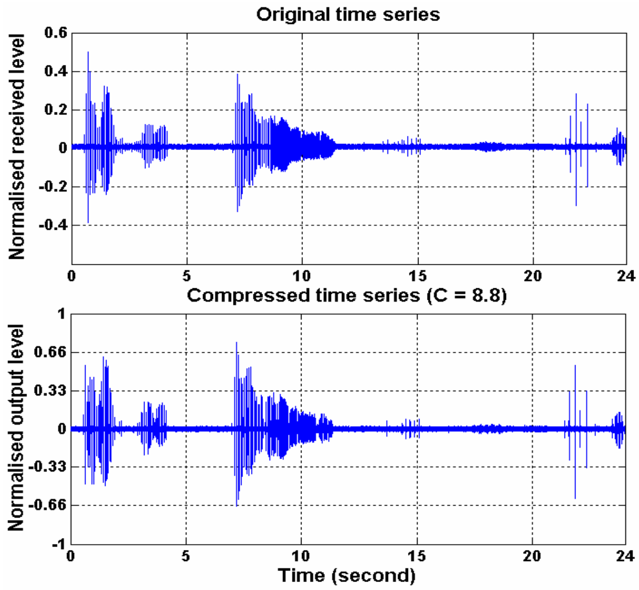


Fig. 6 Comparison between signal shapes of original and compressed dolphin clicks

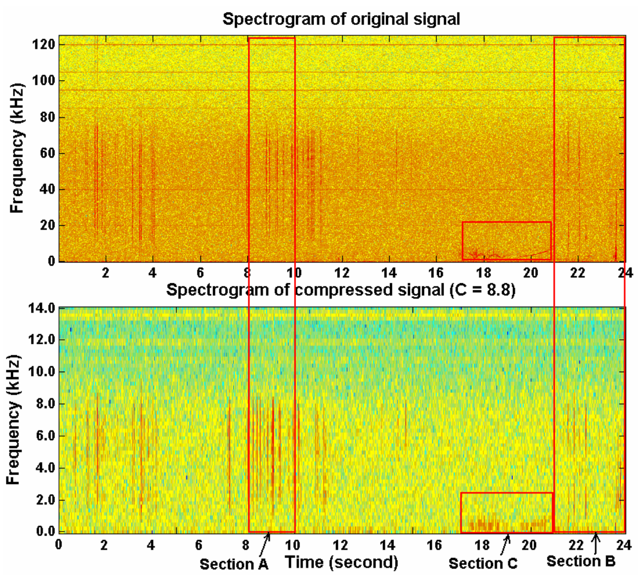


Fig. 7 Comparison between spectrograms of original and compressed ultrasonic dolphin clicks

The second test signal contains a series of broadband dolphin clicks (up to 125 kHz). The bandwidth of the signal is compressed by 8.8 times and the results are given in Fig. 6 and Fig. 7. In this case the compressed time series also follows the amplitude variation of the original signal closely, whereas some clicks were clearly discarded in the

[1] The recording was downloaded from Marine Bioacoustics and Acoustical Oceanography, Università degli Studi di Pavia website. http://www.unipv.it/webcib/cib.html

process of temporal truncation. Nevertheless, the algorithm performed relatively well at preserving the overall nature of the original signal from a perceptual standpoint after playback experiments. For example, the playback of the output shows a click train with the variation of the pulse interval over time resembling the pattern of variations in the original spectrogram.

A section of whistle exists in between the clicks (see Fig. 7, 17-22 *sec* interval and a zoomed display is given in Fig. 9) is successfully mapped to the respective low frequency range. This shows that the algorithm is able to keep the signature of both clicks and whistle that are present in the same time series.

Some part of the original signal in the audio band could be mapped to a lower yet audible spectral range. This would produce duplicated signals (the audible signal from the original time series and the same signal but compressed to lower frequency) that potentially confuses user. This situation can be easily avoided by filtering the audio band before the compression process.

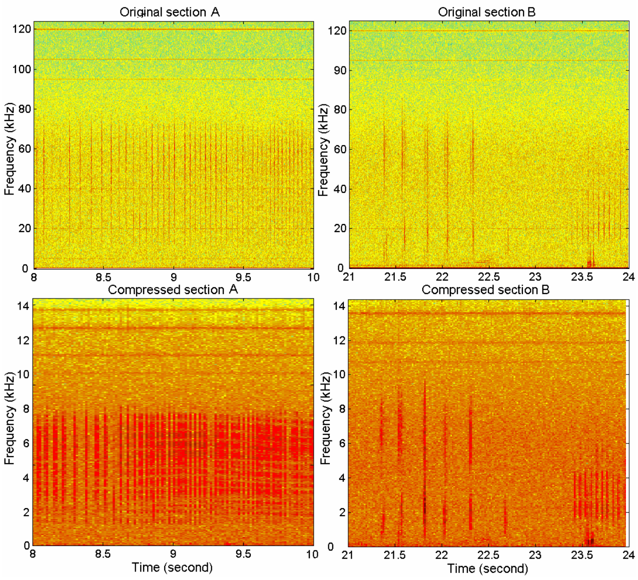


Fig. 8 Blow-ups showing the similarities of click interval variations between original and compressed signals

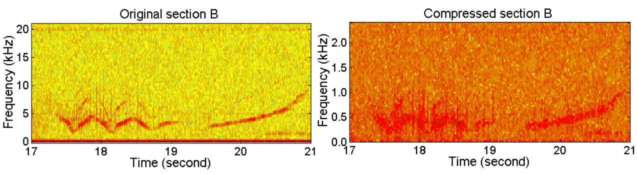


Fig. 9 A section of audible signal from original being compressed into lower but still audible frequencies

## VI. PROTOTYPE OF REAL TIME ABC

An embedded DSP system is currently being prototyped to perform the compression in real time. A simplified block diagram of the prototype system is given in Fig. 10. The system is based on a TI DSP core with a bank of SDRAM. The system takes advantage of the onboard Enhance Direct Memory Access (EDMA) for data transfer, freeing maximum CPU resources for the algorithm. A separate board is used to provide analogue signal conditioning for incoming ultrasound and amplify the outgoing audio signal.

The system is designed for battery operation and the size is kept small so that a researcher can wear it on the waist belt. By changing the input and output transducers, the system could be easily re-configured and used in various in-air and in-water environments.

The system is designed to operate in 2 modes. First operating mode is 'regular chopping'; where the ultrasound is continuously divided into equal blocks and compressed once the system is switched on. The second mode is 'threshold mode'; in this mode, the system will search the time series as the signal comes in for a high-energy region (above a threshold specified by the user). The bandwidth compression is only performed on these blocks.

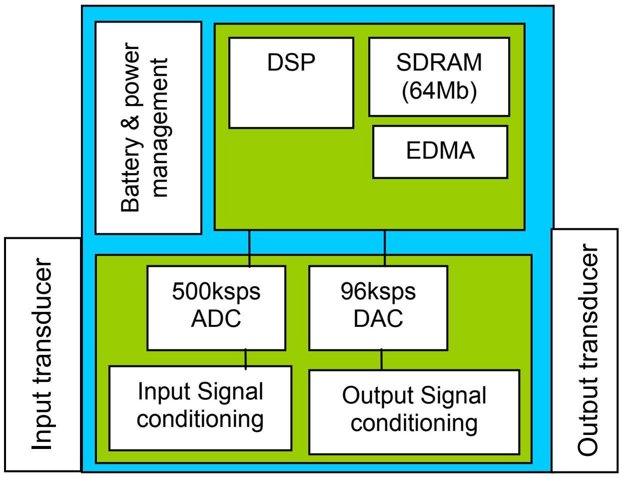


Fig. 10: Simplified block diagram of handheld ABC prototype

## VII. CONCLUSIONS

We have introduced a simple approach to compress the bandwidth of an acoustic signal yet conserving the overall signature of the original signal. As with many other compression methods, this approach loses part of signal information. However instead of discarding large amount of continuous blocks of information in frequency or time domain, we choose to discard the high-resolution frequency structure and maintain the overall audio 'sensation' at the compressed frequency range.

The algorithm is shown to work best when we keep the compression factor just as large as necessary, keep the length of retention windows to around 1 millisecond, and keep the length of smoothing window large. The algorithm has been tested with bio-acoustic signals including long pilot whale whistle and short rapid dolphin clicks. The results show satisfactory compressed outputs with distinctive audio patterns.

A prototype hand held DSP system is being built and currently under test to perform this compression in real time. We anticipate that this algorithm and system will provide the researchers with a new capability to react to ultrasound in real-time, which we expect will offer greatly improved interaction efficiency.


### Acknowledgement

The authors would like to thank Dr. Matthias Hoffmann-Kuhnt and Dr. Adam Pack for making the high frequency dolphin recordings available. This research project is funded by DSTA of Singapore.


### References


[1] K.M. Dudzinski, "Contact behavior and signal exchange among Atlantic spotted dolphins (Stenella frontalis)," *Aquatic Mammals* 24(3): 129-142. 1998.

[2] S.M. Dawson, "Clicks and Communication: The behavioural and social contexts of Hectors dolphin vocalisations," *Ethology* 88(4): 265-276. 1991.

[3] R.M. Seyfarth and D.L. Cheney, "Behavioral mechanisms underlying vocal communication in nonhuman primates," *Animal Learning and Behavior* 25, 249-267. 1997.

[4] M.O. Lammers and W.L. Au, "The broadband social acoustic signaling behavior of spinner and spotted dolphins," *J. Acoust. Soc. Am.* 114 (3), September 2003.

[5] D.R. Wilson and J.F. Hare," Ground squirrel uses ultrasonic alarms," *Nature*, Brief Communications 430, 523. 2004.

[6] K.M., Dudzinski, C.W. Clark, and B. Würsig, "A mobile video/acoustic system for simultaneously recording dolphin behavior and vocalizations underwater." *Aquatic Mammals* 21(3): 187-193. 1995.

[7] D. L. Herzing, ‘‘Vocalizations and associated underwater behaviour of free-ranging Atlantic spotted dolphins, Stenella frontalis and bottlenose dolphin, Tursiops truncatus,’’ *Aqua. Mamm*. 22, 61–79. Herzing, 1996.

[8] P.K. McGregor (Ed.), *Playback and Studies of Animal Communication*. Plenum Press, New York, 1992.

[9] S. Parsons, AM Boonman, and MK Obrist, "Advantages and disadvantages of techniques for transforming and analyzing chiropteran echolocation calls," *Journal of Mammalogy* Vol. 81, no. 4, pp. 927-938. 2000.

[10] S. L. Hopp, M. J. Owren and C.S. Evans, "Animal Acoustic Communication," *Eds. Springer* Verlag, pp. 421 1998.