DETECTION OF CLICKS BASED ON GROUP DELAY

Varvara Kandia[1] **and Yannis Stylianou**[1,2]
1- Institute of Computer Science, FORTH, Crete, Greece
2- Computer Science Dept. University of Crete, Greece

ABSTRACT

In this paper we present a novel approach for the automatic detection of clicks from recordings of beaked whales based on the phase characteristics of minimum phase signals and especially using the group delay function. Group delay is estimated through the and first derivative of the Fourier Transform of a signal. A major advantage of the proposed approach is its robustness against additive noise while it doesn't require the definition of ad-hoc or adaptive thresholds for the detection of clicks. This method works on raw recordings which are usually quite noisy as well as on click enhanced recordings (after band-pass filtering or using operators like the Teager-Kaiser energy operator). Moreover, a click is just detected by searching the positive zero crossings over time of the slope of the phase spectrum. To evaluate the effectiveness of the proposed approach in detecting clicks, a one-minute recording has been manually marked providing a test set of about 320 clicks. Results show that the proposed approach was able to detect 71.37% of the hand labelled clicks within an accuracy of 3 ms.

SOMMAIRE

Dans cet article, nous présentons une nouvelle approche pour la détection automatique de clics sur des enregistrements de baleines à bec exploitant les caractéristiques de signaux à phase minimale notamment via l' utilisation de la fonction de retard de groupe. Le retard de groupe est estimé à partir de la transformée de Fourier d'un signal et de la dérivée de celle-ci. L'approche proposée est robuste vis-à-vis du bruit additif et ne requiert pas la définition ad-hoc ou adaptative de seuils pour la détection de clics. Elle permet de traiter aussi bien des enregistrements bruts fortement bruités que des enregistrements rehaussés (après filtrage passe-bande ou à l'aide d'opérateurs tels que l'opérateur d'énergie de Teager-Kaiser). De plus, un clic est simplement détecté en recherchant un passage par zéro sur la partie croissante de la pente du spectre de phase. Pour évaluer l'efficacité de l'approche proposée à détecter des clics, une minute d'enregistrement a été annotée manuellement, fournissant ainsi un ensemble de test d'environ 320 clics. Les résultats montrent que l'approche proposée parvient à détecter 71.37% des clics marqués manuellement avec une précision de 3 ms.

1. INTRODUCTION

Beaked whales are deep-diving toothed whales and are the least known family of all marine mammals [1]. Two genera of beaked whales, *Ziphius* and *Mesoplodon*, are not as well known as other genera of beaked whales such as *Berardius*. Acoustic monitoring of the sound activity of these animals may help to study their habitats, which is of considerable conservation value since these whales are very difficult to sight. Moreover, there has been a growing concern about the sensitivity of these animals to human-made sounds [2]. Acoustic analysis of the sounds they produce may help in understanding this sensitivity.

Although some whales (i.e., sperm whales) produce sound pulses in the range of human hearing, which is below 20kHz, beaked whales emit short directional ultrasonic clicks (with significant energy above 20 kHz). An analytic report on recordings using acoustic recording tags attached on *Ziphius* and *Mesoplodon* beaked whales may be found in [1] and [3]. Since clicks produced by beaked whales (as well as from other toothed whales) are highly directional, there is a difference in the properties of the signals if they are recorded off or on the acoustic axis of the whale [1]. In the case where hydrophones are used for the recordings, the intensity of the clicks will vary a lot over short periods of time. This makes the detection of clicks harder using energy-based criteria. It modifies the frequency content as well, which complicates the detection problem for the frequency or time-frequency based detectors [4]. Moreover, recordings are usually very noisy which makes the detection task even more difficult. Band-pass filtering is widely used for the enhancement of clicks. However, since the frequency properties of clicks change over time, a time-invariant band pass filter may cancel some of the clicks. More appropriate methods for click enhancement have been suggested in the literature, like in the Rainbow Click detector [5] and the Teager-Kaiser energy operator [6], [7]. To overcome the variability in energy levels of clicks the above approaches need to define adaptive energy thresholds, increasing the complexity of the detection system without significantly improving the detection score.

Therefore, new techniques for the automatic detection and classification of clicks generated from beaked whales are urgently needed to study their behavior and habitat use, and to identify risk factors for exposure of these animals to noise [1]. In this paper time-domain and frequency domain techniques for the automatic detection of the high-frequency clicks of beaked whales will be considered. Although time domain techniques are widely used for detecting clicks, they are not robust in low Signal to Noise Ratio (SNR) conditions as was mentioned earlier. Frequency domain techniques are mainly based on the cross-correlation function defined on the magnitude spectrum of the sounds, ignoring therefore any information provided by the phase spectrum. In this paper we focus on clicks produced by Blainville's beaked whales (*Mesoplodon densirostris*) and we suggest a click detector that combines a time domain technique with frequency domain information based on the slope of the phase spectrum. Specifically, we suggest the use of the Teager-Kaiser energy operator [6] as a click enhancement tool followed by a high resolution group delay function obtained from short-time phase spectra. The group delay function has found important application in numerous signal processing areas, such as speech processing [8]. Clicks are then easily detected by locating the zero crossings of the slope of the phase spectrum (referred to as *phase slope function*) computed as the average of the group delay function. This makes the proposed detector insensitive to variations in sound source level. The proposed detection algorithm performs a frame-by-frame (short-term) analysis. In each analysis window, the slope of the phase spectrum corresponding to the center of the analysis window is computed as the average of the group delay function. Frame (step) size defines the resolution capability of the proposed approach. The algorithm has been tested both on raw recordings of beaked whales and after the application of click-enhancement tools. Results show that the proposed approach is capable of detecting clicks in raw data as well as in pre-processed data (enhanced). For the evaluation of the detector, a one-minute recording was manually labeled providing a set of 317 clicks. Note that in this recording more than one animal is present since the click rate is much higher compared to the mean click rate reported in the literature [1].

The paper is organized as follows. Section 2 describes the time-frequency characteristics of clicks of Mesoplodon beaked whales and tools to improve the SNR in recordings (click enhancement). In Section 3 we present different ways to compute the group delay function and the properties of this function for minimum phase signals are discussed. To motivate the use of the group delay for the detection of clicks, the group delay of synthetic signals similar to a sequence of regular clicks is computed and extensively discussed. Details on the application of group delay for click detection using real click recordings are discussed in Section 4. To evaluate the effectiveness of the proposed approach, clicks have been labeled manually. The database which has been used for the evaluation of the proposed detection system is described in Section 5. A summary of the obtained results and future work concludes the paper.

2. CLICKS OF BEAKED WHALES

Beaked whales are difficult to study and they are mostly known from strandings. They are deep-diving animals, they echolocate on prey [1] and they react to human-made sounds. In [1] two genera of beaked whales, *Ziphius* and *Mesoplodon* have been tagged making orientation and sound recordings. The tagged whales started clicking at an average depth of 400-500 m. Both species produce regular clicks with an inter-click-interval (ICI) of about 0.4s for *Ziphius* and between 0.2 and 0.4s for *Mesoplodon*. Regular clicks usually terminate with a buzz sound (rapid increase in click rate, 250 clicks per second [1]). The average duration of the clicks were measured at 175 μs and 250 μs for *Ziphius* and *Mesoplodon*, respectively. For both species, the energy of their sounds is mostly distributed at high frequencies, i.e., in the 30kHz-40kHz range.

2.1 Time-frequency information

In Fig. 1(a) 13 seconds of a recording from beaked whales (*Mesoplodon densirostris*) are depicted[1]. Sounds were digitized at a sample rate of 96 kHz, with 24-bit resolution. From this figure, it is not easy to detect clicks by inspecting the time domain signal. In Fig. 1(b) the time-frequency distribution of the signal is displayed, computed via the Short-Time Fourier transform using a Hanning window of 1000 samples (10.4 ms) with an overlap of 500 samples, and a frequency resolution of 2048 bins. In Fig. 1(b), looking at frequency bands above 20 kHz, some wideband signals may be detected (for instance, around 4, 5, and 8 seconds), indicated the presence of clicks. The high frequency content of the clicks is expected after the results presented in [1]. It is worth noting that by comparing the two figures, it is obvious that the detection of clicks is easier in the time-frequency domain. This is the motivation for using such a representation for the detection of clicks with software products such as Ishmael [4].

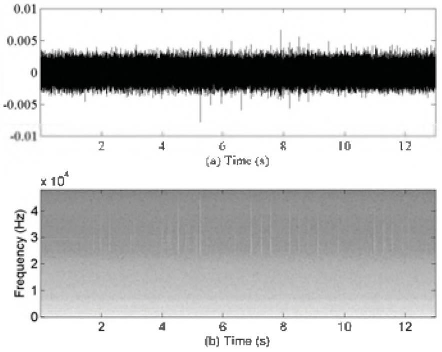

Figure 1. Typical recording from a beaked whale (a) time-domain signal, sampled at 96 kHz, 24 bits, and (b) short-time Fourier transform (2048 frequency bins, Hanning window of 1000 samples, with 50% overlap).

2.2 Click-enhancement

The signal depicted in Fig. 1(a) is very noisy and it is not easy (if not impossible) to detect any click by visual inspection. Therefore, a click enhancement tool could

---

[1] This recording is part of the recordings made available by the Naval Undersea Warfare Center (NUWC) Marine Mammal Monitoring on Navy Undersea Ranges (M3R) program.

possibly reveal the clicks and improve the SNR. For this purpose, we will use 2 enhancement tools: one is based on the Teager-Kaiser energy operator [6] and the other is based on modulation and downsampling.

Teager-Kaiser energy operator

For a discrete time signal x[n], it is shown in [9] that the Teager-Kaiser (TK) energy operator is given by

$$\Psi\left[x(n)\right] = x^2(n) - x(n+1)x(n-1) \qquad (1)$$

where n denotes the sample number. An important property of the TK energy operator in (1) is that it is nearly *instantaneous* given that only three contiguous samples are required in the computation of the output at each time instant. More details on the TK operator as applied to click sounds may be found in [6, 7].

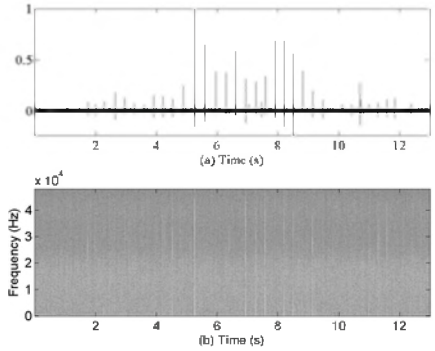

Figure 2. After applying the Teager-Kaiser operator (a) the output from the Teager-Kaiser (TK) operator and (b) the short-time Fourier Transform of the TK output using the same setup as in Fig.1.

Fig. 2(a) shows the results from the application of the TK operator on the original recording, while Fig. 2(b) displays the associated Short-Time Fourier transform. Contrary to Section 2.1, the time-domain signal now carries more information about the time occurrence of clicks than its time-frequency representation. Using an adaptive energy threshold, detection of most of the clicks would be possible. Note that the sampling frequency of the signal has not been changed.

Modulation and down-sampling

Low SNR makes it difficult to detect clicks. Moreover, since most of the energy of clicks is distributed to frequencies higher than 20 kHz, it is also hard to detect them aurally. Therefore, clicks cannot also be detected aurally. Based on the frequency characteristics of the emitted clicks we suggest to modulate the amplitude of the signal and then down-sample it appropriately. Let's denote the original signal by x[n]. Then the output signal, y[n] from the above operations is given by:

$$\upsilon\left[n\right] = x\left[n\right]\cos\left(\pi n\right)$$
$$w\left[n\right] = \upsilon\left[\downarrow 2\right]$$
$$y\left[n\right] = w\left[n\right]\cos\left(\pi n\right)$$

where $\downarrow 2$ denotes the downsampling operation by two after applying an anti-aliasing lowpass filter. The last modulation in the above equations is required in order to re-establish the order of the frequency information. The time domain signal (y[n]) and the time-frequency distribution of the signal are depicted in Fig.3. As for the Teager-Kaiser operator, the clicks are clearly seen in the time domain signal. Although some of the narrow band signals are also present in the time-frequency distribution, not all of the clicks are visible. Therefore, again in this case, one would prefer the time domain representation to detect the clicks.

Although, after the application of the above enhancement tools, the "click structure" was revealed in the previous examples, it is still difficult to detect a great number of clicks because of the variability of the click intensities. We recall that the intensity is a function of the position of the whale relative to a hydrophone. Since the whale is moving the intensity changes quickly because of the high directional characteristic of clicks. Taking into account the possibility that other beaked whales are present in the area and may also emit clicks at different distances from the hydrophone, the click intensity can vary quickly over a short period. In this case, a sophisticated time-adaptive system of thresholds should be used for click detection. To overcome this, we suggest to use the slope of the phase spectrum, computed as the average of the group delay function, for click detections. This will make the click detector insensitive to the plethora of different click intensities.

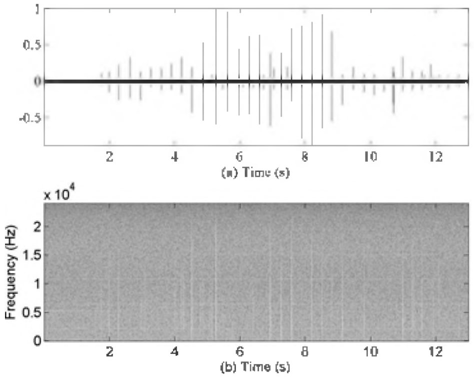

Figure 3. After modulation and down-sampling (a) time-domain signal, sampled at 48 kHz, 16 bits, and (b) short-time Fourier Transform (1024 frequency bins, Hanning window of 500 samples, shifted by 250 samples).

3. GROUP DELAY FUNCTIONS

3.1 Motivation

Consider a delayed unit sample sequence $x\left[n\right] = \delta\left[n - n_0\right]$ and its Fourier Transform $X(\omega) = e^{-j\omega n_0}$. The group delay is defined as [10]:

$$\tau(\omega) = -\frac{d\phi(\omega)}{d\omega} \qquad (2)$$

so the group delay for the delayed unit sample sequence is $\tau(\omega) = n_0 \ \forall \omega$, since the phase spectrum of the signal is $\phi(\omega) = -\omega n_0$. The average over ω of $\tau(\omega)$ provides n_0 which corresponds to the negative of the slope of the phase spectrum for this specific signal and to the delay of the unit sample sequence. An example of a delayed unit sample sequence with $n_0 = 200$ samples as well as the associated group delay function are depicted in Fig. 4(a) and (b), respectively. In this example the Fourier Transform has been computed considering the center of analysis window to be at $n = 0$. When the window center is moved to the right (closer to the instant $n = n_0$), the slope of the phase spectrum increases (the average of the group delay function decreases) reflecting the distance between the center of the analysis window and the position of the impulse. When the center of the analysis window is at $n = n_0$ then the slope is zero. Continuing moving the analysis window to the right the slope will continue to increase (while the average of the group delay will decrease). In this way, the slope of the phase spectrum is a function of n. Note that the location of the zero-crossing of this function will provide the position of the non-zero value of the unit sample sequence independently of the amplitude value of the impulse. Filtering the unit sample sequence by a minimum phase system, results in a minimum phase signal with the same delay as the input unit sample sequence. The group delay function will still provide information about this delay value as well as information about the poles of the minimum phase system. In Fig. 4(c), (d) the output of the minimum phase signal and the associated group delay are depicted. The slope function will have a similar behaviour to this described earlier for the unit sample sequence.

By creating a periodic sequence of minimum phase signals as the one displayed in Fig. 4(c), a sequence similar to a train of regular clicks may be obtained. Defining an analysis window of length proportional to the period of the sequence (it will be referred to as *long window*), a frame-by frame analysis is performed. In each frame the slope of the phase spectrum of the windowed signal is computing and it is associated at the center of analysis window. By setting the analysis step size at one sample (moving the analysis window by one sample at a time), the obtained phase slope function (signal) has the same time resolution as the original recording. The window length may have a duration shorter than the period of the signal (it will be referred to as *short window*). In Fig. 5(a) the periodic sequence of the minimum phase signal is displayed along with the phase slope function using long (dashed line) and short (dash-dotted line) window. As it was expected based on the description provided before, the positive zero crossings of the slope function provide the position of the "clicks". Of course, the detection of clicks using a simple energy criterion will provide the same detection score as the proposed approach, in this example. In Fig. 5(b), the same sequence of "clicks" is repeated but now the energy of the minimum phase signals linearly decreases as time increases. In the same figure the phase slope function is also displayed using, as in Fig. 5(a), the same types of lines for long and short analysis window. It is obvious that a simple energy criterion will not work as well as before and an adaptive energy criterion should be used. Instead, using the slope of the phase function, the position of "clicks' are still easily detected. This is expected since the phase information is not related to the total energy of a signal but rather to the distribution of the signal energy over time.

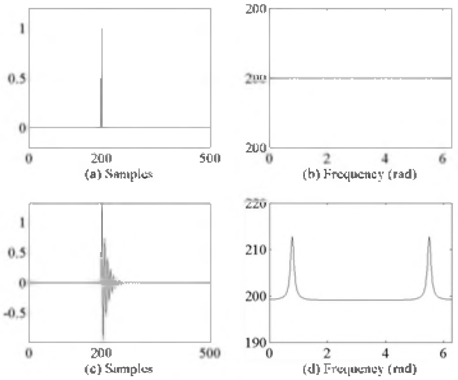

Figure 4. (a) A delayed by 200 samples unit sample sequence. (b) The group delay function of the signal in (a). (c) A minimum phase signal with an oscillation at $\pi/4$. (d) The group delay function of the signal in (c).

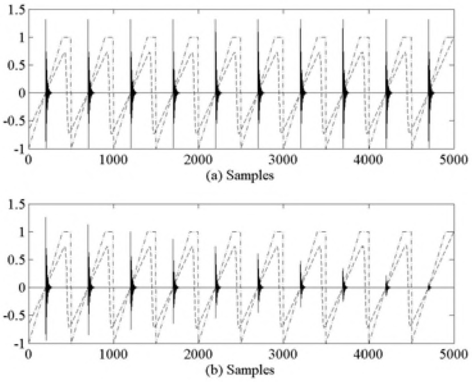

Figure 5. (a) A sequence of impulses of constant amplitude and the associated phase slope function using long (dashed line) and short (dash-dotted line) window (b) A sequence of impulses with linearly time varying amplitudes and the associated phase slope function using long (dashed line) and short (dash-dotted line) window.

3.2 Computing group delay

To compute the group delay of a signal or the average slope of the phase spectrum we need to compute the unwrapped phase spectrum. This is necessary because phase is computed modulo 2π and any attempt to compute the slope using wrapped phase data will produce erroneous results. Usually phase unwrapping is performed by adding appropriate integer multiples of 2π to the principal phase values, as to remove discontinuity (jumps of 2π radians) in the phase curve. Unfortunately, phase unwrapping is not always successful. Therefore, we suggest to compute the slope of the phase function through an alternative to Eq. (2) computation of the group delay [10]:

$$\tau(\omega)=\frac{X_R(\omega)\Upsilon_R(\omega)+X_I(\omega)\Upsilon_I(\omega)}{|X(\omega)|^2} \quad (3)$$

where

$$X(\omega)=X_R(\omega)+jX_I(\omega)$$
$$\Upsilon(\omega)=\Upsilon_R(\omega)+j\Upsilon_I(\omega)$$

are the Fourier Transforms of $x[n]$ and $nx[n]$, respectively. Using Eq. (3) we avoid the computation of the unwrapped phase. The phase slope is then computed as the negative of the average of the group delay function.

4. DETECTION ALGORITHM

Clicks from beaked whales are highly directional and of very short duration. They can therefore be seen as realizations of impulse responses of minimum phase systems. For the application of the phase slope function to the detection of clicks, we set the length of the analysis window as a function of the average inter-click interval. According to Johnson et al. [1] the average inter-click interval for Blainville's beaked whales is about 0.3s. So for the experiments shown below, we used a hanning window of 0.5s (long window). In this section, the example of the recording shown in Fig.1 will again be considered. Comparing the original recording and its enhanced versions with the ideal train of pulses presented earlier, it is expected that the first step before the computation of the phase slope will be the application of an enhancement tool. In the upper panel of Fig.6 the output from the Teager-Kaiser operator on the original recording is repeated, while in the middle panel the associated phase slope function is depicted. By detecting the positive zero crossings of the phase slope function, the location of clicks is obtained. This is shown in the lower panel of Fig.6, where a unit sample sequence is generated according to the positive zero crossings. It is worth to note the high correlation between the train of clicks and the unit sample sequence. Similar results are obtained if the modulated and downsampled version of the original recording is used. The associated results are shown in Fig.7.

Finally, we have applied the phase slope function on the original recording without the application of any enhancement tool or any other pre-processing step. To our surprise, the structure of clicks is clearly revealed! We believe that this result merits further investigation. Results are depicted in Fig.8. It is worth to note the similarity of results with and without the use of enhancement tools. By comparing closely the detection results in these three cases, we found that there are some differences in detecting the clicks which is more noticeable in the noisy areas of the signal. When the SNR is high[2] (for instance between 5 and 8 s, in Fig.6), the obtained results are very similar.

To evaluate a click detector, hand labeled data are required. Part of the signal shown in Fig.7 and specifically between the 5th and 8th s, is depicted in Fig.9(a) along with a series of hand labels shown as little triangular marks. It is worth to note the presence of very low intensity clicks in addition to the clicks with relatively high intensity. Labels for the low intensity clicks are also shown. We assume that these clicks are recordings by a conspecific made off the acoustic axis of the whale. To also detect these low intensity clicks using the phase slope function, the window length has to be short enough. For this example the window length was set to 0.1s. In Fig.9(a) the phase slope function is also displayed by a dashed line. It is worth to note the high correlation between the labels and the positive zero-crossings of the phase slope function. If we were to use a longer window, for instance a window of duration 0.5s, then only the high intensity clicks would have been detected. In Fig. 9(b) the detection of clicks is indicated by a sequence of unit samples. For this example, the mean absolute error between the manually labeled and the automatic detected click instances is 1.1ms. Finally, there is deletion of one click at around 6.7s and one extra click detected at around 7.8s. The mean absolute error in this case is low, however, there are some cases where the detection is not so accurate. The degree of accuracy mainly depends on the SNR. Thus, despite that the structure of clicks can be revealed under very low SNR conditions using the phase slope function, comparisons with hand labeled data shows that using an enhancement tool improves the accuracy. As we have seen, the modulation and the downsampling process improves the SNR of the original recordings. Besides, this is the signal that a human will use to mark the clicks. We suggest to improve further the SNR of that signal by applying the Teager-Kaiser operator on the output after the downsampling. For the short example discussing above, this shows an improvement in accuracy. Indeed, using the TK operator the mean absolute error was decreased to 0.03 ms. However the number of deleted or inserted clicks remained the same as before (one click is deleted and one is inserted). The proposed click detection system is shown in Fig.10.

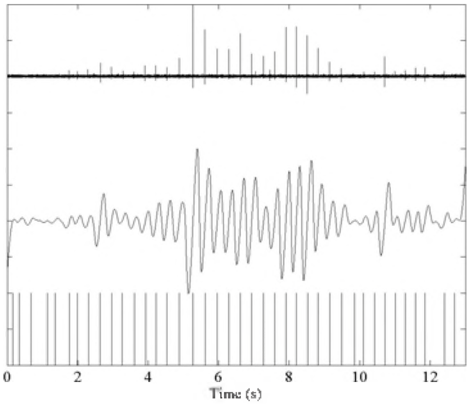

Figure 6. Upper panel: the output from the Teager-Kaiser operator. Middle panel: the associated phase slope function. Lower panel: detection of clicks based on the positive zero crossings of the phase slope function.

---

[2] Although SNR is not so meaningful for the original recording, since none of the clicks are easily detected by visual inspection.

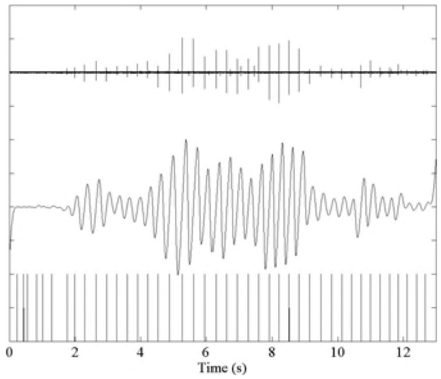

Figure 7. Upper panel: the output from the modulation and down-sampling operations. Middle panel: the associated phase slope function. Lower panel: detection of clicks based on the positive zero crossings of the phase slope function.

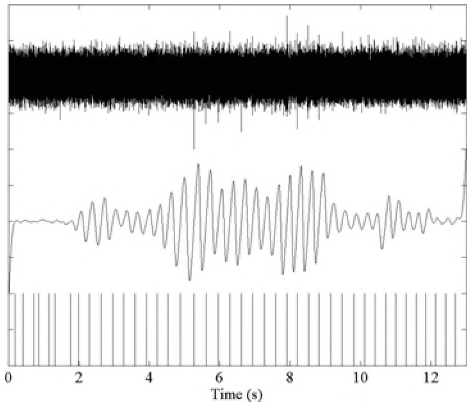

Figure 8. Upper panel: the original recording. Middle panel: the associated phase slope function. Lower panel: detection of clicks based on the positive zero crossings of the phase slope function.

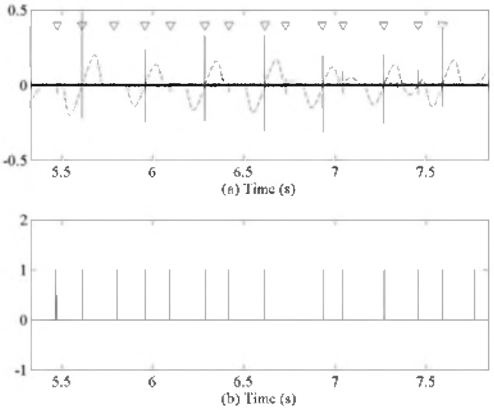

Figure 9. (a) the original recording after modulation and downsampling (solid line), the manual labels (triangle), and the phase slope function (dashed line). (b) Positive zero crossings of the phase slope function indicated by a sequence of unit samples.

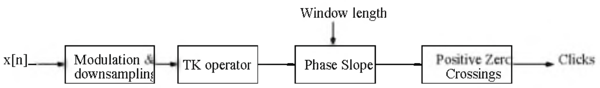

Figure 10. Block diagram of the proposed click detection system.

5. APPLICATION

The proposed click detection system has been evaluated on the training dataset provided by the organizers of the 3rd International Workshop on Detection and Localization of Marine Mammals using Passive Acoustics (Boston, MA, USA, 2007) and were recorded by the Naval Undersea Warfare Center. More specifically, we used recordings of Blainville's beaked whales recorded at a sample rate of 96 kHz, with 24 bits accuracy, from Set1, Alesis number 2, hydrophone H18 at AUTEC (one minute) to manually labeled and from Set3 and 4, Alesis 3 and 6, hydrophones H27 and H76 for visually control the click detection results.

5.1 Database

One minute of recording[3] was manually labeled providing in total 317 clicks. These labeled clicks will be referred to as *ds1* dataset. The recording contained only regular clicks. However, the presence of more than one animal was evident in this recording because of the density of clicks and the variability in their intensity. From a visual inspection of the original recording, it was not possible to detect any of these clicks. The initial recording was modulated and downsampled to 48kHz and it was listened through a closed type headphone in a quite office room. Marking was facilitated by using the Sound Forge software. To improve the accuracy of the marking the playback speed of the sound was considerably lowered in some cases (i.e., to about 2kHz). This facilitates the auditory and visual inspection of fast moving acoustic events. It is worth noting that some of the clicks were not easily detectable visually, while only a trained person could hear them. The decision was made to remove some of the clicks, creating therefore a second dataset for the evaluation of the system. This set contained 248 clicks and it will be referred to as *ds2* dataset.

In addition to the manually labeled dataset, the recordings mentioned above from hydrophones H27 and H76, were analyzed and inspected visually. The phase slope function and the automatically detected clicks (positive zero crossings of the phase slope function) were displayed along with the TK output of the modulated downsampled signals. In this way, the correlation between the positive zero crossings of the phase slope function and the clicks could be checked relatively quickly.

5.2 Results

The proposed system was tested on the manually labeled data as well as other recordings contained in the provided training dataset. For all tests, a Hanning analysis window of duration 0.15s was used. To speed up the computation of the slope phase function, a step size of 30 samples was used. Undetermined values of the slope function were computed by linear interpolation. A click was assumed to be detected if the absolute difference between the time-instances of the manually and the automatically detected

---

[3] Filename: Set1-A2-092405-H18-0000-0030-1008-1038loc_0000-0100min.wav

clicks was within 3ms. If this difference was between 3 ms and 20 ms, the click was assumed to be missed (deletion), while a difference over 20 ms indicated a click insertion. For the evaluation of the system two criteria were used; the detection rate (referred to as *Det*) and the corrective rate (referred to as *Corr*). The detection rate is defined as:

$$Det = \frac{Number\ of\ clicks\ correctly\ detected}{Total} \times 100$$

where *Total* is the total number of manually labeled clicks, and the corrective rate is defined as:

$$Corr = \frac{Total - Deleted - Inserted}{Total} \times 100$$

where *Deleted* referred to as the number of clicks that were considered to be missed (deleted) and *Inserted* refers to the number of extra clicks that have been inserted by the proposed system. It is worth noting that the phase slope function shows occasionally oscillations of very low amplitude about zero which are mostly associated to the noise. On the contrary, for clicks we observe high-amplitude oscillations about zero. Therefore, by subtracting a constant (dc component) from the phase slope function, most of the erroneous clicks associated with noise or very low intensity clicks embedded into noise were eliminated. Such a subtraction was done in order to use the same phase slope function for the two different sets of labels: *ds1* and *ds2*. For *ds1* and *ds2*, the constant was set to 0.1% and 1% of the maximum value of phase slope function, respectively, eliminating 24 erroneous clicks for *ds1* and 7 erroneous clicks for *ds2*.

Table 1 summarizes the detection results for the two datasets, *ds1* and *ds2* (we recall that *ds2* is a subset of *ds1*). As expected, results are better for the second dataset. The number of clicks detected by the system was comparable to the number of manually detected clicks; 321 for *ds1* and 253 for *ds2*. The detection score as well as the corrective rate were mostly affected by the missed clicks. For instance for *ds1*, by increasing the lower threshold (tolerance) to 6ms, the detection score is 85.64% and the corrective rate is 78.12% (compare to 64.03% and 63.72%, respectively). Most of the missed clicks were clicks closely located and for their discrimination a shorter analysis window was needed.

The system was also tested on the training test data mentioned above which contained recordings from Blainville's beaked whales. By visual inspection it has been found that the positive zero-crossings of the phase slope function corresponded to clicks in most cases while some erroneous clicks were introduced from occasional oscillations about zero of the slope phase function as discussed above. Again, by subtracting a constant (dc component) from the phase slope function most of these erroneous clicks were eliminated. Also some clicks were missed because the analysis window was not short enough for their detection (closely spaced clicks).

6. SUMMARY AND CONCLUSIONS

In this paper we present a new technique for detecting clicks from beaked whales based on group delay. More specifically we use the slope of the phase spectrum which is computed as the average value of the group delay function of the input signal. The approach is insensitive to the intensity of clicks and it is robust against very low SNR conditions.

	Clicks	Corr (%)	Det (%)
ds1	317	63.72	64.03
ds2	248	71.37	72.98

Table 1. Detection results using a tolerance of 3ms.

Combined with the Teager-Kaiser operator, a click detection system was developed and evaluated using recordings from Blainville's beaked whales (*Mesoplodon densirostris*). The proposed system was tested on recordings with manually labeled clicks as well as on unlabeled recordings. Results show the effectiveness of the proposed system in detecting clicks. Mainly, the only design parameter of the system is the length of the analysis window and dc offset. Window length controls the details of the detection. We plan to use the proposed system on clicks from sperm whales and compare its performance with frequency domain and energy based click detectors. Alternative ways in computing the slope of the phase spectrum will be considered.

7. ACKNOWLEDGMENTS

The authors would like to thank the Naval Undersea Warfare Center (NUWC) for making the data used in this paper publicly available, and the organizers of the 3nd International Workshop on Detection and Localization of Marine Mammals using Passive Acoustics (Boston, MA, USA, 2007), for providing access to these data.

8. REFERENCES

1. M. Johnson, P.T. Madsen, W. M. X. Zimmer, N. A. de Soto and P. L. Tyack. “Beaked whales echolocate on prey.” Proc. Royal Soc. Biology Letters, 271:383-386, 2004.
2. Frantzis. “Does acoustic testing strand animals?” Nature, 329(29), 1998.
3. P. T. Madsen, M. Johnson, N. A. de Soto, W. M. X. Zimmer and P. L. Tyack. “Biosonar performance of foraging beaked whales (*Mesoplodon densirostris*).” The Journal of Experimental Biology, 208:181-194, 2005.
4. D. K. Mellinger. “Ishmael 1.0 Users Guide”, NOAA, NOAA/PMEL/OERD, 2115 SE OSU Drive, Newport, OR 97365-5258, 2001. Technical Memorandum OAR PMEL-120.
5. D. Gillespie. “An acoustic survey for sperm whales in the Southern Ocean sanctuary conducted from the R/V Aurora Australis”, Rep. Int. Whal. Comm. 47:897-908, 1997.
6. V. Kandia and Y. Stylianou. “Detection of creak clicks of sperm whales in low SNR conditions.” In CD Proc. IEEE Oceans, Brest, France, 2005.
7. V. Kandia and Y. Stylianou. “Detection of sperm whale clicks based on the Teager-Kaiser energy operator.” Applied Acoustics, 67(11-12):1144-1163, 2006.
8. R. Smits and B. Yegnanarayana. “Determination of instants of significant excitation in speech using group delay function.” IEEE Trans. on Speech and Audio Processing, 3(5):325-333, 1995.
9. J. F. Kaiser. “On a simple algorithm to calculate the ‘Energy’ of a signal.” In Proc. IEEE ICASSP, pages 381-384. Albuquerque, NM, USA, 1990.
10. A. V. Oppenheim, R. W. Schafer and J. R. Buck. “Discrete-Time Signal Processing.” Prentice Hall, 1998.