Aquatic Mammals 2005, *31*(4), 463-467, DOI 10.1578/AM.31.4.2005.463

An Attractive Alternative for Sperm Whale Click Detection Using the Wavelet Transform in Comparison to the Fourier Spectrogram

Maciej Lopatka,[1, 2] Olivier Adam,[1] Christophe Laplanche,[1] Jan Zarzycki,[2] and Jean-François Motsch[1]

[1]*Laboratoire d'Images, Signaux et Systèmes Intelligents (Lissi-iSnS), University Paris 12, France*
[2]*Signal Theory Section, Institute of Telecommunications and Acoustics, Wroclaw, Poland*

Abstract

Although many mathematical and signal-processing tools exist, detection of sperm whales based on their sound recordings proves somewhat difficult. This paper presents some of the advantages of the wavelet transform and the spectrogram in analyzing sperm whale clicks. The coefficients of the wavelet transform and the short-time Fourier transform are used to provide a representation of the intrinsic characteristics of the sound emissions of the sperm whale. For detection, we propose a new parameter—Short-Time Windowed Energy—that characterizes the particular shape in the time-frequency/scale domain of sperm whale clicks. This paper illustrates the resistance to noise of this parameter. In addition, thanks to new processors, this algorithm, which was once lengthy in calculation time, can be integrated easily in a real-time system.

Key Words: sperm whale, *Physeter*, click, detection, wavelet transform, Fourier transform

Introduction

Visual observation is a widely used technique for analyzing the behavior of marine mammals; however, this requires that the animals be present at the water surface—a trait that is not typical of the behavior of the sperm whale (*Physeter macrocephalus*), which tends to swim in very deep waters (> 2 km), rarely rising to the surface (< 10% of the duration of a complete dive).

Various methods can be used to track the sperm whale. Electronic sensors can be fixed on the animal's back (Zimmer et al., 2003), and an active acoustic sonar can be used; however, it is possible that these approaches could perturb the animal and consequently alter its behaviour. An alternative method consists of passively recording the signals emitted by the sperm whale. Much information can be attained regarding its detection, identification, and trajectory. These are some of the subjects of interest to cetologists (Watkins & Schevill, 1977; Whitehead & Weilgart, 1990; Goold & Jones, 1995; Goold, 1996; André & Kamminga, 1999; Jaquet et al., 2001; Zimmer et al., 2003).

One of our research areas focused on time-frequency analysis based on parametric models such as the generalized Schur algorithm (Zarzycki, 2004). This paper contributes towards further research regarding detection of the sperm whale. In it, we introduced the definition of a new parameter—Short-Time Windowed Energy (*STWE*)—which results from the comparative study of two different approaches used to deduce the presence of the animal: (1) the spectrogram and (2) the wavelet transform (Gordon, 1991; Mellinger & Clark; 1993, Dougherty, 1999; Lopatka, 2002). The performances of these approaches are compared (Adam et al., 2005), with special attention given to their resistance to noise.

Time-Frequency Representation of the Clicks Emitted by the Sperm Whale

For the purpose of this study, signals were recorded onboard the CIRCé research ship, *Elsa*, during the scientific experiment organized in May 2003 by the Centre d'Etudes Biologiques de Chizé (CEBC-CNRS, France) in the Strait of Gibraltar. At that time, we observed three sperm whales. We used HC2000 Vinci-Technology hydrophones, with a 80Hz – 22 kHz band pass. The data were sampled at 44.1 kHz and coded by 16 bits. We used a 2nd order highpass filter (1 kHz) to attenuate the ambient noise coming largely from ships nearby.

Signals emitted by sperm whales have particular characteristics (*cf.* Figure 1a). They are nonstationary, brief, transitory, and have a large spectral band (*cf.* Figure 1b). These characteristics can obstruct the use of some estimators, particularly in the quest to obtain a satisfying time-frequency representation.

Spectrogram of the Fourier Transform

The advantage of this approach is the ease of its implementation in software. The short-time Fourier transform is obtained through fast calculation. In our studies, we used the well-known split radix method to reduce the calculation time (Vetterli & Duhamel, 1989). Moreover, its time-frequency representation is very simple to interpret (*cf.* Figure 1b). Nevertheless, this technique has some drawbacks when used with sperm whale clicks. First, its transitory and brief signal makes the choice of the weighted time window difficult (Lopatka, 2002). Fixed a priori, its length will be systemically suboptimal. The compromise was to use either temporal resolution or frequency resolution, but not both simultaneously. For our purposes of detection, we chose to use *temporal* resolution. We used Kaiser's time window with length 1.43 ms, $\beta = 5$, and each sample overlapping to obtain the finest time resolution (Haykin, 1994). Increasing the value of parameter β widens the main lobe and decreases the amplitude of side-lobes (with $\beta = 5$, attenuation = 54 dB).

Please note that for all figures, white represents minimum and black represents maximum values.

For the detection problem, we selected the representative frequency range and defined the parameter, which is calculated as follows:

$$IE_{STFT}[f, lT_e] = \sum_{f=f1}^{f2} \left| b_w[f, lT_e] \right|^2 \tag{1}$$

with b_w the short-time Fourier transform coefficients, f the frequency, $f1$ and $f2$ the frequency range of the click's spectrum, T_e the sampling period, and l the time resolution.

Wavelet Transform

The wavelet transform is based on functions limited both in time and frequency (Mallat, 1988; Daubechies, 1994; Strang & Nguyen, 1996; Bialasiewicz, 2000) in place of the complex exponential kernel of the Fourier transform.

For this study, we used the Daubechies15 wavelets chosen empirically from our dataset. These wavelets are orthogonal and without analytic expression. The continuous wavelet transform of the signal $x(t)$ is calculated as follows:

$$X_w(a,b) = \int_{-\infty}^{+\infty} x(t) \frac{1}{\sqrt{a}} \Psi^*\left(\frac{t-b}{a}\right) dt \tag{2}$$

with $X_w(.)$ the wavelet coefficients, $a > 0$ the dilatation coefficient, and $b \in \Re$ the shifting term.

In our particular case, the main advantage of the wavelet transform was that the length of the temporal window is adapted to the analysed frequency.

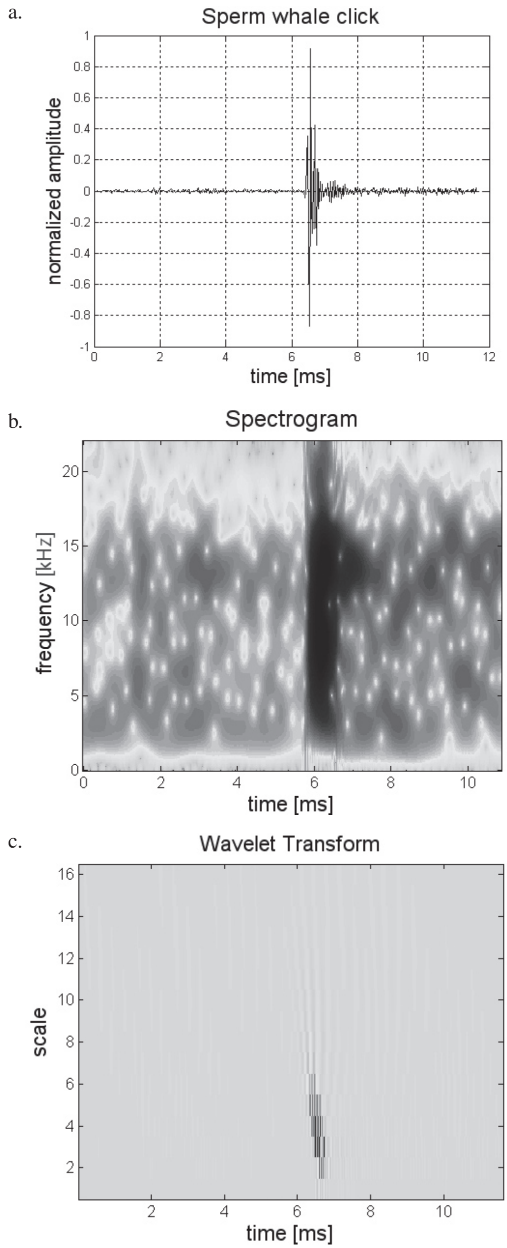

Figure 1. (a) Waveform of a typical sperm whale click; (b) spectogram of this click; and (c) wavelet transform of this click

In addition, the wavelet transform provided a time-scale representation (*cf.* Figure 1c). This "fanning out" shape is characteristic of brief rectangular signals such as those emitted by the sperm whale.

Additionally, we can extract pertinent information from the coefficients of the wavelet transform. These parameters can be used to substantially reduce noise (Daubechies, 1994). Most importantly, these parameters can be used for deciphering the exact time the click emitted by the sperm whale clicks was recorded. To isolate this moment in the continuously temporal recording, we defined *STWE* for the continuous wavelet transform as

$$STWE_{WT}\left[s_k, lT_e\right] = \sum_{k=k_1}^{k_2} c_w^2\left[s_k, lT_e\right] \quad (3)$$

with c_w the wavelet transform coefficients, s_k the scale, k_1 and k_2 the scale range of the click's wavelet transform, T_e the sampling period, and l defining the time resolution.

Detection Algorithm

$STWE_{STFT}$ and $STWE_{WT}$ were used for detecting the presence of sperm whale clicks in continuously recorded signals. For both cases, . . .

- the resulting *STWE* signal was divided on windows of 4,000 samples (for the frequency sampling 44.1kHz, it is possible to have no more than several clicks in one window).
- each window was analyzed for searching maximum relative values bigger than a given threshold or higher than some value proportional to the mean value of *STWE*.
- the InterClick Interval (ICI) was the time between two successive clicks, calculated from the maximum instantaneous pressure peak of one click to the next; it characterizes the rhythm. ICI was used to limit the maximum number of clicks in each window. This parameter avoids taking into account direct echoes of the click (from the sea surface or from the sea bottom).

The apex of the *STWE* curves (*cf.* Figure 2b and 2c) illustrated the exact time when the sperm whale click was recorded. The width represents the duration of the click.

Figure 2 shows the relative *STWE* amplitude that we used for detecting the sperm whale click presence. Note that the $STWE_{WT}$ shows close results (compared with $STWE_{STFT}$) while the two clicks vary for a constant signal-to-noise ratio.

Results and Discussion

The statistical variability of the clicks is taken into account by considering the successive clicks in the recorded signals. Sperm whales emit repetitive clicks for the duration of their dive (except right before surfacing); therefore, it is easy to calculate the ICI during the recorded signal. We should note that the spectrum of sounds emitted by the sperm whale varies in relation to the individual whale, its behavior and sex (different types of clicks), and its distance from the hydrophone (Gordon, 1991; Dougherty, 1999; Jaquet et al., 2001; Lopatka, 2002). For this reason, we normalize (in a shift short-time window) these ranges of scales and frequencies of successive clicks to minimize the effect of variability, especially for the detection of clicks emitted by the same sperm whale.

After presenting the advantages and drawbacks of these two methods, we will now compare their resistance to noise. We added white noise to the recordings, and the SNR varied between 15 dB and -10 dB. We then compared the deterioration of this time-frequency representation in both the Fourier estimator and the wavelet transform when the SNR diminishes (*cf.* Table 1). The time frequency obtained for signals with SNR = 15 dB served as a

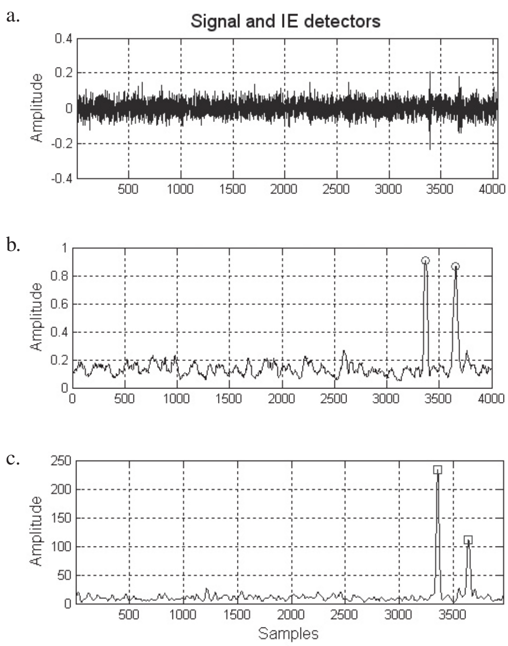

Figure 2. (a) Recorded signal with additional gaussian white noise (SNR = -10dB), containing two clicks; (b) detection with $STWE_{WT}$; and (c) detection with $STWE_{STFT}$

reference. Then, we compared the results obtained when the SNR decreased.

Table 1. Deterioration (%) of the time-frequency representation with the SNR

SNR (dB)	15	10	5	0	-5	-10
STFT	100	85	67	48	28	9
WT	100	90	78	69	56	46

The spectrogram based on the Fourier transform is less resistant to noise (*cf.* Table 1); in the presence of significant noise, it is difficult to venture conclusions from the results. Because of the fixed resolution on time-frequency representation, the detection was more difficult, and the positions of the apex are less precise. In addition, the use of a threshold presents some drawbacks, particularly when the signal-to-noise ratio weakens.

With the wavelet transform, we filtered the noise contained in this signal by choosing to use only the first four coefficients (high frequencies), representing signal information. The $STWE_{WT}$ gave the same results as obtained without the presence of noise. The maximum of the $STWE_{WT}$ curve remains the same, and we can continue to detect the presence of the sperm whale.

We performed the Receiver Operating Characteristic (ROC) curves for our two detectors to show their performances (Figure 3). Note that the $STWE_{WT}$ was more robust, giving the better results as $STWE_{STFT}$. The best result obtained was a hit probability of 97% and the probability of false alarm of 4.5%.

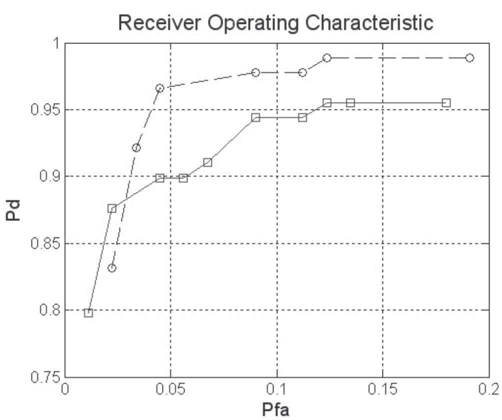

Figure 3. ROC curve of the two detectors ($STWE_{WT}$- circles and $STWE_{STFT}$ - squares)

Conclusions

This paper presents two different methods used to compare the results obtained using the sperm whale signals recorded in the Strait of Gibraltar.

We present the advantages of each method using the wavelet transform as the results obtained prove promising for the detection of the sperm whale click in a continuous recording. In so doing, we used two parameters from the coefficients of the Fourier transform and the wavelet transform.

These results showed that we can choose definitively the $STWE_{WT}$ for contributing to the automatic sperm whale detection. Wavelet analysis provides an attractive alternative sperm whale click detection approach that is more robust to noise than traditional STFFT detection processing.

Acknowledgments

Thanks are given to the CEBC-CNRS (France), CIRCé (Spain), MAREE (France), and Association DIRAC (France). Many thanks to the anonymous reviewers for their contribution in improving the clarity and the rigorousness of this paper. Thanks to Jeanette Thomas for refining the English.

Literature Cited

Adam, O., Lopatka, M., Laplanche, C., & Motsch, J. F. (2005). Sperm whale signal analysis: Comparison using the autoregressive model and the wavelet transform (International Conference on Soft Computing and Intelligent Systems [SCIS]). *Enformatika, 1*, 188-195. (ISBN 975-98458-3-0)

André, M., & Kamminga, C. (1999). Rhythmic dimension in sperm whale echolocation click trains: Identification and communication? In *European research on cetaceans 13* (pp. 226-234). Valencia, Spain: European Cetacean Society.

Bialasiewicz, J. T. (2000). *Falki i aproksymacje*. Warszawa: Wydawnictwa Naukowo-Techniczne.

Daubechies, I. (1994). Ten lectures on wavelets. *CBMS, SIAM, 61*, 194-202.

Dougherty, A. M. (1999). *Acoustical identification of individual sperm whales*. Master's of Science thesis, University of Washington, Seattle, Washington, USA.

Goold, J. C. (1996). Signal processing techniques for acoustic measurement of sperm whale body lengths. *Journal of the Acoustic Society of America*, *100*(5), 3431-3441.

Goold, C., & Jones, S. (1995). Time and frequency domain characteristic of sperm whale clicks. *Journal of the Acoustic Society of America*, *98*(3), 1279-1291.

Gordon, J. C. D. (1991). Evaluation of a method for determining the length of sperm whales from their vocalizations. *Journal of Zoology (London)*, *224*, 301-314.

Haykin, S. (1994). *Signal processing*. New York: IEEE Press.

Jaquet, N., Dawson, S., & Douglas, L. (2001). Vocal behaviour of male sperm whales: Why do they click? *Journal of the Acoustic Society of America*, *109*(1), 2254-2259.

Lopatka, M. (2002). *Reconnaissance de signatures acoustiques pour la distinction d'individus dans un groupe de cachalots* (iSnS Report, University of Paris 12). Paris, France.

Mallat, S. (1988). *A wavelet tour of signal processing*. New York: Academic Press.

Mellinger, D. K., & Clark, C. W. (1993). A method for filtering bioacoustic triansients by spectrogram image convolution. *Oceans*, *3*, 122-127.

Strang, G., & Nguyen, T. (1996). *Wavelets and filter banks*. Wellesley, MA: Wellesley-Cambridge Press.

Vetterli, M., & Duhamel, P. (1989). Split-radix algorithms for length-p^m DFT's. *IEEE Transactions on Acoustic, Speech, and Signal Processing*, *37*(1), 57-64.

Watkins, W., & Schevill, E. (1977). Sperm whale codas. *Journal of the Acoustic Society of America, 62*(6), 1485-1490.

Whitehead, H., & Weilgart, L. (1990). Click rates from sperm whales. *Journal of the Acoustic Society of America*, *87*(4), 1798-1806.

Zarzycki, J. (2004). Multidimensional nonlinear Schur parameterization of nonGaussian stochastic signals. Part two: Generalized Schur algorithm. *MDSSP Journal*, *15*, 243-275.

Zimmer, W., Johnson, M. P., D'Amico, A., & Tyack, P. L. (2003). Combining data from a multisensor tag and passive sonar to determine the diving behaviour of a sperm whale. *IEEE Journal of Oceanic Engineering*, *28*(1), 13-28.