Acoustics'08 Paris
June 29-July 4, 2008
www.acoustics08-paris.org

Accurate estimation of the duration of tonal signals emitted by marine mammals

Nicolas Sentenac

Thales Underwater Systems, 525, route des dolines, BP 157, 06903 Sophia Antipolis, France
nicolas.sentenac@fr.thalesgroup.com

Passive marine mammals detection is of great importance nowadays. The goal of this article is to present a tool that can be used to accurately measure the duration of tonal emissions from marine mammals.
Marine mammals emissions can be split into two different kinds. The first one is narrow band pattern emission, and the second one is echo-localization clicks that sound like impulsive broadband signals. This article focalizes on the former.
Narrow band signals can be detected in the time/frequency space thanks to the use of the spectrogram and thanks to a contrast criterion derived from the one proposed by Buffington. This article will introduce the expression of a lower bound of the expectation of the criterion for noisy signal.
As this lower bound is a function of signal to noise ratio, it can be used to threshold the criterion with the desired signal to noise ratio value. When the criterion is computed along time on a sliding window, narrow band signals with a higher snr than the thresholding ones are detected as soon as they appear and until they vanish, giving an accurate estimation of the duration of such signals.

1 Introduction

Passive marine mammals detection is of great importance nowadays. The goal of this article is to present a tool that can be used to accurately measure the duration of tonal emissions from marine mammals, helping to classify them. The classification part will not be treated here.

Marine mammals emissions can be split into two different kinds. The first one is tonal emission, which can be defined by their narrow band frequency pattern, and the second one is echo-localization clicks that sound like impulsive broadband signals. This article focalizes on the former.

Narrow band signals can be detected in the time/frequency space thanks to the use of the spectrogram and thanks to a contrast criterion derived from the one proposed by Buffington in 1974 [1]. The expression of this criterion for signal-only and noise-only is straightforward. This article will introduce the expression of a lower bound of the expectation of the criterion when signal and noise and mixed.

As this lower bound is a function of signal to noise ratio, it can be used to threshold the criterion with the desired signal to noise ratio value. When the criterion is computed along time on a sliding window, narrow band signals with a higher snr than the thresholding ones are detected as soon as they appear and until they vanish, giving an accurate estimation of the duration of such signals.

2 Definition of Buffington's criterion

Given a signal whose Power Spectral Density (PSD) is X, one of the definition of Buffington's criterion computed on the frequency band B is:

$$C = \frac{\int_B X(f)^2}{\left(\int_B X(f)\right)^2} B \tag{1}$$

When the signal is sampled at F_s, X is estimated on N samples from the signal:

$$\hat{X}_k \approx X\left(k\frac{F_s}{N} \le f < (k+1)\frac{F_s}{N}\right), \forall k \in [0, N/2-1] \tag{2}$$

So C can be rewritten with the M estimates of X in the band B, dealing with a frequency resolution equal to F_s/N :

$$C = \frac{\frac{F_s}{N}\sum_{k\in B}\hat{X}_k^2}{\left(\frac{F_s}{N}\sum_{k\in B}\hat{X}_k\right)^2} M\frac{F_s}{N} = \frac{\sum_{k\in B}\hat{X}_k^2}{\left(\sum_{k\in B}\hat{X}_k\right)^2} M \tag{3}$$

For greater convenience M will be dropped in the latter equation. It does not make a big difference in the demonstration but it makes the curves clearer and emphasizes the impact of a small value of M.

$$C = \frac{\sum_{k\in B}\hat{X}_k^2}{\left(\sum_{k\in B}\hat{X}_k\right)^2} \tag{4}$$

3 Criterion for white noise only

It is commonly known that for a white noise of variance σ_o^2, the values of its PSD (estimated by a single spectrum) are distributed according to a χ^2 law with two degrees of freedom.

$$E(\hat{X}) = 2\sigma_o^2/N \tag{5}$$

The second order moment of a χ^2 distribution with $2M$ degrees of freedom equals to $2(2M)+(2M)^2$, so:

$$E\left(\left(\sum_{k\in B}\hat{X}_k\right)^2\right) = \alpha^2\left(2(2M)+(2M)^2\right) \tag{6}$$

Also, the second order moment of a χ^2 law with 2 degrees of freedom is 8, so:

$$E\left(\sum_{k\in B}\hat{X}_k^2\right) = \sum_{k\in B}E\left(\hat{X}_k^2\right) = 8\alpha^2 M \tag{7}$$

Finally, the criterion expresses itself by:

$$E(C_{BB}) \approx \frac{E\left(\sum_{k\in B}\hat{X}_k^2\right)}{E\left(\left(\sum_{k\in B}\hat{X}_k\right)^2\right)} = \frac{2}{1+M} \tag{8}$$

This expression is needed in the next paragraph.

4 Criterion for noisy narrow-band signal

Given a narrow-band signal which temporal amplitude is A, hidden in a Gaussian white noise of variance σ_o^2.

The denominator of the criterion expresses the sum of the frequency bins containing the energy of both signal and noise.

$$E\left(\left(\sum_{k\in B}\hat{X}_k\right)^2\right)=\left(M\frac{2\sigma_o^2}{N}+\frac{A^2}{2}\right)^2 \tag{9}$$

The numerator contains the sum of the squared energy of the signal. Considering the spectral spread due to temporal windowing (Hann window is used here), the energy of the signal is spread over four frequency bins. The lower bound of the numerator sum can be computed. As shown on Fig. 1, the estimate of the central frequency can be anywhere between two multiples of the frequency resolution associated to the DSP computation.

The sum of squared levels for different shift is given on Fig.2:

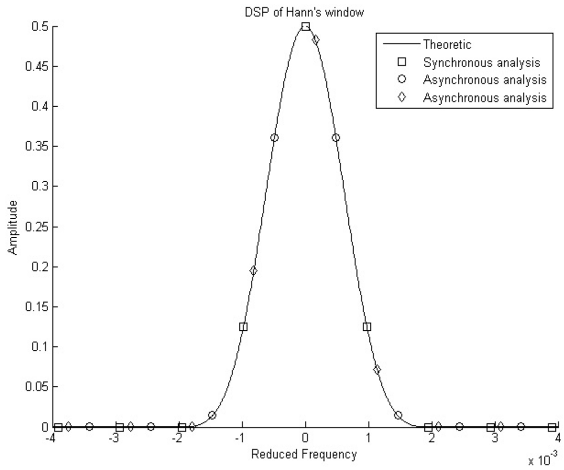

Fig. 1 Measured DSP of Hann's window depending on frequency shift

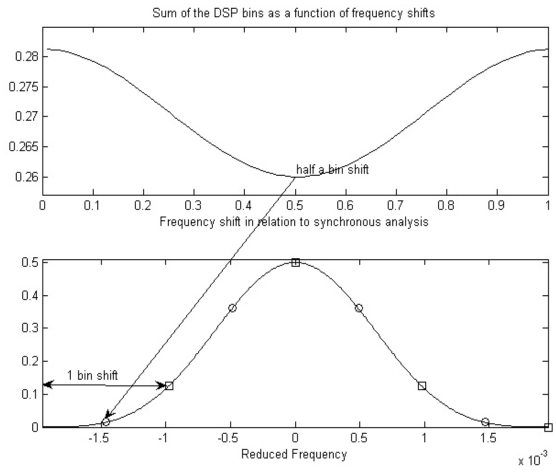

Fig. 2 Sum of main DSP levels as a function of frequency shift

As forecast, the lower bound of the summed main level of the DSP is obtained when the frequency of the sine is at the exact mean of two consecutive frequency bins.

The squared Fourier transform of Hann's window is given by:

$$\|FT_{hann}(\upsilon)\|^2=\frac{1}{8}(1-\cos(2\pi\upsilon N))\left|\frac{1}{\pi\upsilon N\left((\upsilon N)^2-1\right)}\right|^2 \tag{10}$$

Knowing that the gain of Hann's window in relation to noise is 2/3, a lower bound of the energy of a half-bin shifted bin is obtained:

$$\begin{aligned} E&=2\left\|FT_{hann}\left(\frac{1}{2N}\right)\right\|^2\frac{2A^2}{3}\\ E&=\frac{64A^2}{27\pi^2}\end{aligned} \tag{11}$$

However, tests have been led and showed that in the case of a transition between a noise only signal and a noisy sine signal, the sine wave might be present only during a fraction of the FFT (beginning of the sine signal) and it results in a lower value of the DSP. To take into account this case, it is advised to use another lower bound of the energy. It corresponds to the presence of the sine wave during two thirds of the length of the FFT:

$$\tilde{E}=\frac{3}{4}E=\frac{16A^2}{9\pi^2} \tag{12}$$

Taking the two adjacent main bins of the DSP *i* and *j*:

$$\begin{aligned} \sum_{k\in B}\hat{X}_k^2 &= \sum_{k\in B/\{i,j\}}b_k^{\,2}+(S_i+b_i)^2+(S_j+b_j)^2\\ \sum_{k\in B}\hat{X}_k^2 &\geq (M-2)\frac{8\sigma_o^4}{N^2}+2\left(\frac{2\sigma_o^2}{N}+\frac{16A^2}{9\pi^2}\right)^2\end{aligned} \tag{13}$$

The lower bound, which can be seen as a threshold of the criterion, is given by:

$$E(C_{BE})\geq\frac{8(M-2)\sigma_o^4/N^2+2\left(2\sigma_o^2/N+16A^2/9\pi^2\right)^2}{\left(2M\sigma_o^2/N+A^2/2\right)^2} \tag{14}$$

Now, trying to discriminate a noisy sine wave from its background noise with a given signal to noise ratio *snr* (before FFT), the later expression rewrites:

$$snr=10\log_{10}\left(A^2/2\sigma_o^2\right)$$

$$E(C_{BE})\geq\frac{8(M-2)\sigma_o^4/N^2+2\left(2\sigma_o^2/N+16\cdot2\sigma_o^2\cdot10^{snr/10}/9\pi^2\right)^2}{\left(2M\sigma_o^2/N+2\sigma_o^2\cdot10^{snr/10}/2\right)^2}$$

$$E(C_{BE})\geq\frac{2(M-2)+2\left(1+16N\cdot10^{snr/10}/9\pi^2\right)^2}{\left(M+N/2\cdot10^{snr/10}\right)^2} \tag{15}$$

To illustrate this expression, let us consider a short sine wave (present between 1 and 2) mixed with noise. Fig.3 shows both the value of the criterion and the value of the threshold corresponding to the exact *snr* of the sine wave along time. A good correspondence between the threshold and signal to noise ratio can be seen, particularly at the lowest *snr*.

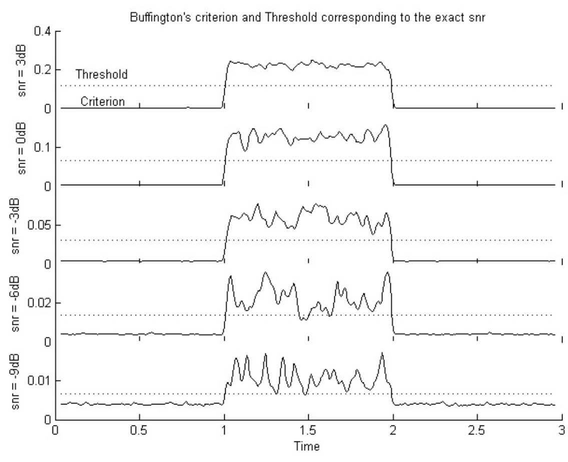

Fig. 3 Threshold as a function of the *snr*

This fact allows a thresholding depending on a minimum *snr* to detect.

For example, a threshold computed for an *snr* of -3dB allows detecting a signal whose *snr* is -3dB (signal in the dotted box on Fig.4) whereas a signal whose *snr* is -6dB (signal in the dashed box) is not.

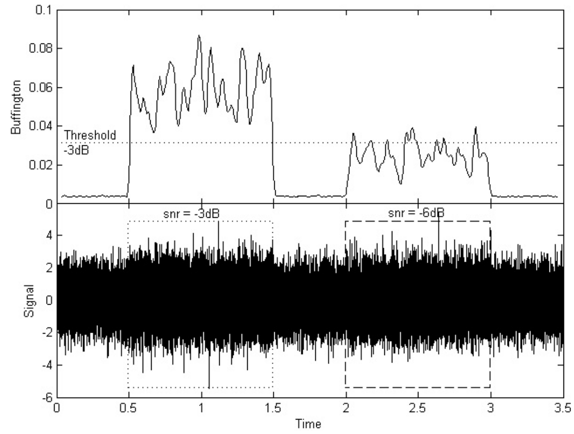

Fig. 4 Example of rejection by the threshold

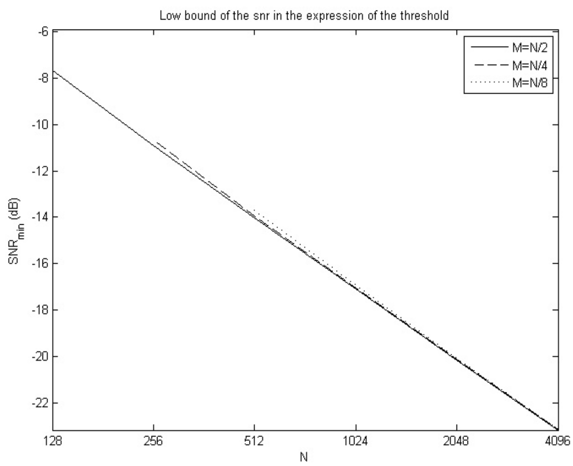

Fig. 5 Minimum value of the *snr* used in the threshold as a function of *N* and *M*

A signal whose *snr* is lower than the *snr* used in the threshold can make the criterion higher than the threshold, but it happens once in a while and is not very annoying.

Of course, the threshold must be higher than the criterion value for noise only (cf. Eq.(8)). Depending on the analysis parameters, this constraint gives a lower bound for the *snr* to be used in the expression of the threshold:

$$rsb > 10\log\left(\begin{array}{l}\dfrac{18\pi^2}{N\left(1024(M+1)-81\pi^4\right)}\times\ldots\\ \left(\begin{array}{l}9M\pi^2-32(M+1)+\ldots\\ \ldots\sqrt{(M+1)\left(1024(2+M)-576M\pi^2+81(M-1)\pi^4\right)}\end{array}\right)\end{array}\right) \quad (16)$$

The value of this lower bound varies very slowly with the value of *M*. Fig.5 shows the lower bound of the tresholding *snr* as a function of *N,* the number a point is the DSP computation.

Let us add that the number *M* of frequency bins used in the criterion must be large enough. When *M* drops, the criterion value tends to be higher than the corresponding threshold, reaching a common value whatever the signal *snr* is. The variance of the criterion changes also with the value of *M*.

Fig.6 shows for *N*=1024 and for signal *snr* of -6, -3 and 0dB, the corresponding threshold and a zone where the criterion value is when computed on 500 randomly generated noisy signals. As these zones are overlaid, they have been displayed partially transparent. Dashed lines mark the outer bounds of the zones.

As said before, three characteristics of the threshold can be observed. First, the threshold is really a lower bound at low *snr*. Second, the criterion value tends to be the same for every signal *snr* when *M* is small. And last, the variance of the criterion value varies with *M*.

The fact that the criterion value tends to a common value for any *snr* when *M* decreases lessens the threshold accuracy. Indeed, for a small value of *M* the criterion value, computed on a noisy signal, can rise higher than the threshold corresponding to a really higher thresholding *snr*.

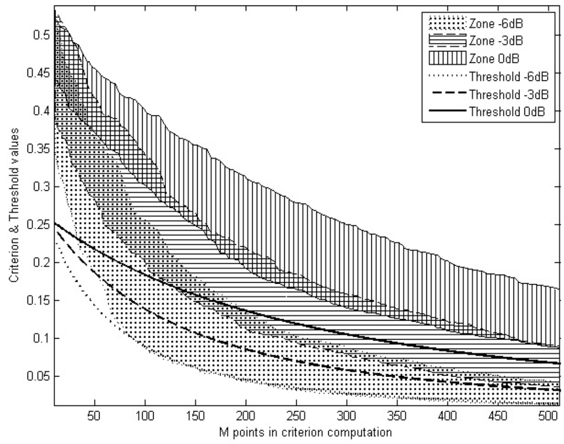

Fig. 6 Zones occupied by the criterion value for different noisy signals and corresponding thresholds

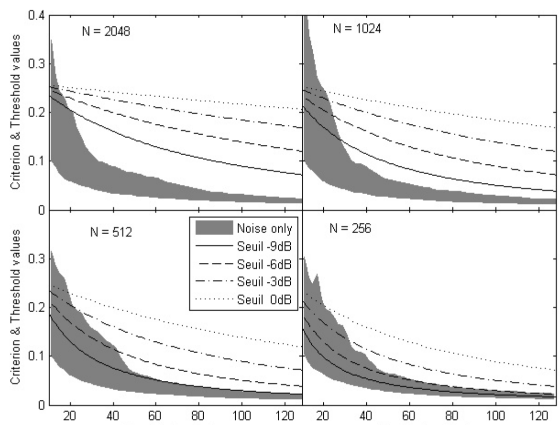

Fig. 7 Criterion value for white noise as a function of M

Fig.7 shows, for *N*={2048, 1024, 512, 256}, the zone occupied by the criterion computed upon 500 different white noises and the threshold for different *snr*. The main observation is that false alarms may occur more often for small values of *M*. Computing the criterion with about fifty points seems to be a lower bound when trying to detect low *snr* signals without raising the false alarm occurrences.

5 Examples of application

Given two different signals emitted by Killer whales, spectrograms are computed upon 2048 windowed points with strong overlap as shown on Fig. 8 and 9:

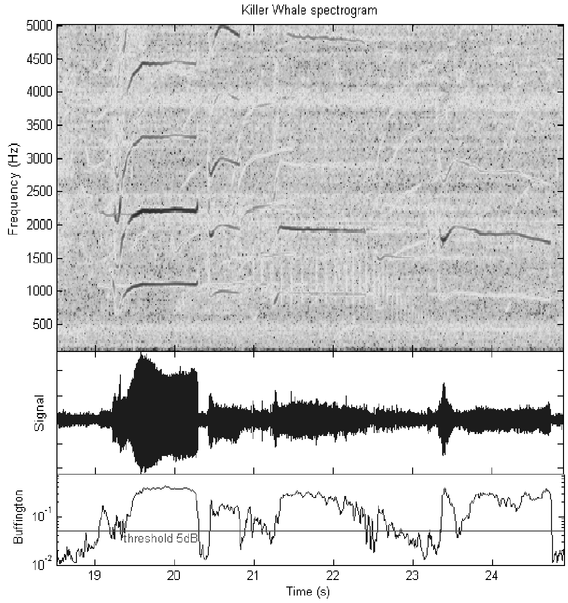

Fig. 8 Killer whale Spectrogram, criterion and threshold

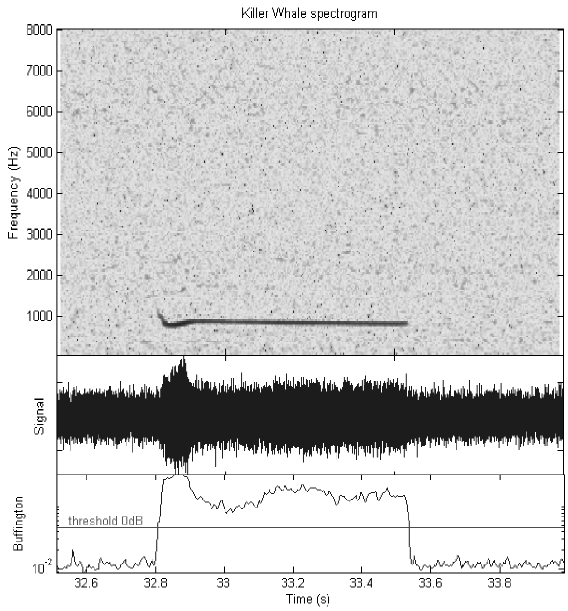

Fig. 9 Killer whale Spectrogram, criterion and threshold

These figures shows that even with multiple harmonics, the threshold is robust, avoids numerous false alarms and gives a good indication of the presence of tonal signals.

5 Conclusion

This article focused on the accurate estimation of the duration of tonal signals emitted by marine mammals. A spectral contrast criterion has been used to do so and a threshold has been introduced to allow a good separation of noise-only signals and noisy narrow-band signals. Moreover, this threshold can be tuned with a true signal to noise ratio value, permitting to reject most of the signals that do not reach this *snr* value and detect all signals above this *snr* value. The ability to set a threshold according to a desired *snr* value is more convenient than setting it empirically.

Acknowledgments

This work has been partially supported by French Service des Programmes Navals (DGA)

References

[1] R.A. Muller, A. Buffington "Real time correction of atmospherically degraded telescope images through image sharpening", *J. Opt. Soc. Am.*64 (1974)