Estimation of the number of whale individuals based on click sounds of selected whale species

Carlos de Obaldía Pastor[1], Udo Zölzer[1]

[1] *Helmut Schmidt University, 22043 Hamburg, Germany, Email: deobaldia@hsu-hh.de*

Abstract

In this work an automatic abundance estimation method for the passive acoustic monitoring (PAM) of cetaceans is proposed. A system for the segmentation and extraction of transients in underwater monaural recordings is presented which is applied for the detection of cetacean clicks. Vocalizations of the Harbour Porpoise, Blainville's Beaked Whale, and Cuvier's Beaked Whale are analyzed in different frequency bands, and a transient enhancement method is used to reduce tonal noise sources. Differentiated click trains are then identified by comparing each click successively using a correspondence analysis for the determination of the number of active individuals in the signal.

Introduction

Passive acoustic monitoring (PAM) techniques and methods are used to model and analyze sound waves in a passive way. By monitoring underwater sounds we can look for particular cetacean vocalizations in order to determine their presence, the quantity of individuals in a particular area, and their trajectories and movements. Measures for the conservation and protection of marine mammals and their ecosystems can be thus enhanced by enabling a system for estimation of the population density of a particular species in a region [1].

Cetaceans typically communicate using frequency-modulated pulses (clicks) and tonal contours (sweeps), among other types of calls. Clicks are used for echolocation, foraging as well as for social purposes. These clicks can as well contain information of the active individual which can be studied for a better understanding of cetacean behavior [2, 3].

The study in [2] introduced a click detection, segmentation and extraction system to determine the number of individuals of sperm whales in monaural recordings. In this work, the algorithm is expanded to generalize the concept for click detection and segmentation of the Harbour Porpoise (Phocoena phocoena), the Cuvier's Beaked Whale (Ziphius cavirostris) and the Blainville's Beaked Whale (Mesoplodon densirostris). A transient detection algorithm is applied to underwater communication signals of several species of the cetacea family. A filter bank is used to detect click species in different bands and clicks are grouped according to their apparent click train. Figure 1 depicts a block diagram for the proposed PAM system.

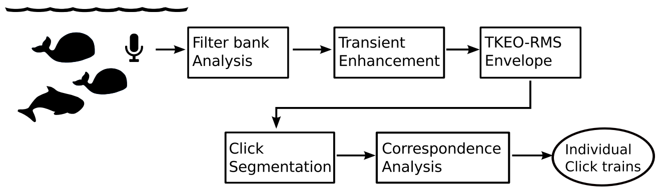

Figure 1: Passive acoustic monitoring framework for individual identification. The signal is first divided in different sub-bands, each of which are processed for click detection and segmentation using the TKEO-RMS detector after enhancing transients. Clicks are then segmented and compared to each other to obtain individual click trains.

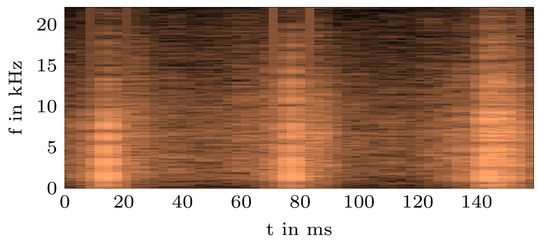

Figure 2: Time-frequency representation of a signal containing low-frequency clicks of a Harbour Porpoise. Clicks appear as vertical ridges in the spectrogram

Segmentation and Extraction

Click patterns appear in the spectrogram as high energy vertical regions, which can be broad-band, appearing on a large frequency range, or narrow-band, where just particular frequencies are present at the event. These clicks thus exhibit a transient-like characteristic in the spectral and time domains. Figure 2 shows the frequency domain representation of a recording excerpt where low frequency clicks of a Harbour Porpoise are present.

Click Detection

Since clicks can occur in the time domain signal as abrupt rises in energy bounded to short periods of time, an energy detector can be employed to detect such rises. Segmentation of the click segments is modified from the algorithm presented in [2] to account for the generalized case. For a particular species, clicks can be finely extracted by analyzing the peaks of the Teager-Kaiser energy operator (TKEO) contour [2]. However, since clicks have a particular structure depending on the species emitting the sound, in this case just the peak of the RMS envelope is taken in consideration for the extraction. The click segments are then delimited between the onset and decay of the root-mean-square (RMS) envelope peaks. The

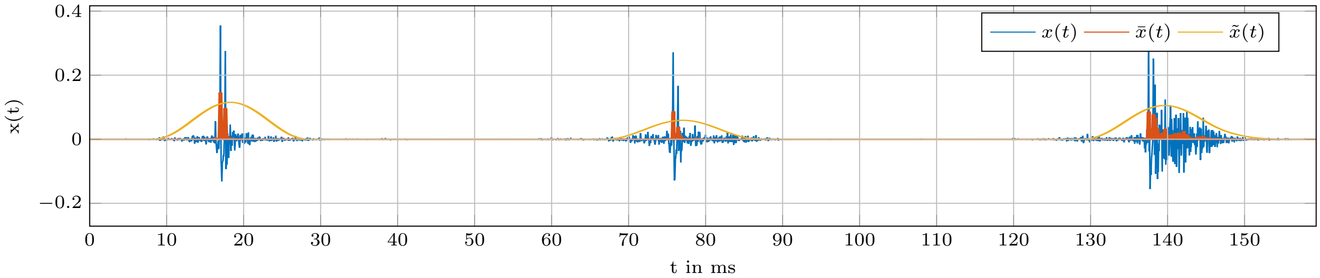

Figure 3: TKEO-RMS extraction over LF clicks of a clean Harbour Porpoise recording. The blue signal is the original recording, the orange is the TKEO output and the yellow line is the RMS envelope. Clicks are segmented between minima surrounding each peak of the RMS envelope.

TKEO is applied on the discrete time signal $x(n)$ as:

$$\bar{x}(n) = \frac{1}{2}\{x(n)^2 - x(n+1)\cdot x(n-1)\}. \tag{1}$$

To define click delimiters, the RMS envelope over a window of 20 ms of the extracted energy contour is calculated by

$$\tilde{x}(n) = \sqrt{\frac{1}{L_{RMS}} \sum_{n-\frac{L_{RMS}}{2}}^{n+\frac{L_{RMS}}{2}} \bar{x}(n)^2}, \tag{2}$$

where $L_{RMS} = f_s \cdot 20\,\mathrm{ms}$ and f_s is the sampling frequency for the signal. The clicks are then segmented between two consecutive minimums of the RMS envelope as it is depicted in Fig. 3.

Transient Enhancement

The underwater environment suffers from an ever increasing problem of noise pollution. Most of these noises arise from, or correspond to anthropogenic sources like ships, commercial sonars, depth finders, drilling platforms and active underwater communication systems [4]. These noise sources are difficult to model since their behavior is not time invariant due to the underwater transmission channel characteristics [5] which can particularly affect Cetaceans in their behavioral patterns as they cannot perform tasks like locating prey and engaging in social conduct [6].

If the signal-to-noise ratio of the clicks decay, the energy of the noise can mask time domain transients thus effectively hiding the clicks from detection. To overcome this a harmonic/percussive separation algorithm as in [7] is used for enhancing apparent vertical transients in the spectrogram.

The spectrogram $X(b,k)$ of an arbitrary signal $x(n)$ where b denotes the current frequency bin and k a time frame can be reformulated as the sum of two parts

$$X(b,k) = H(b,k) + P(b,k) \tag{3}$$

where $P(b,k)$ represents the transient components in the spectrogram and $H(b,k)$ the harmonics. The *transient* spectrogram, $P(b,k)$ is extracted by multiplying the spectrogram with a mask $M_p(b,k)$

$$P(b,k) = X(b,k) \cdot M_p(b,k), \tag{4}$$

which is the result of soft-masking the transient and harmonic extracted components such that

$$M_p(b,k) = \frac{\tilde{P}(b,k)^2}{\tilde{H}(b,k)^2 + \tilde{P}(b,k)^2}, \tag{5}$$

being $\tilde{P}(b,k)$ the median filtered spectrogram along the frequency and $\tilde{H}(b,k)$ the median filtered spectrogram along time, which are computed by

$$\tilde{H}(b,k) = \mathcal{M}(X(b), l_{harm})\ \forall k \quad \text{and} \tag{6}$$

$$\tilde{P}(b,k) = \mathcal{M}(X(k), l_{perc})\ \forall b, \tag{7}$$

where $\mathcal{M}(x,l)$ is the median filter operator and l is the filter length. For this work, a filter length of $l_{harm} = 17$ and $l_{perc} = 50$ is used. The output of the algorithm can be used for source separation of different shapes in the frequency domain [7, 8]. However, in this work just the *percussive* parts, $P(b,k)$, are used for enhancement of the clicking regions. The Fourier transform of Eq. 4 is then used for detection in the time domain as it is shown in Fig. 4.

Multiband processing

Before extracting clicks and segmenting them in the incoming signal, the signal is filtered in adjacent bands. By band-limiting the signal in different frequency regions, clicks can be discovered irrespectively of the influence of high energy noises in other frequency regions. Transients in the signal can then be in turn further emphasized to enhance active click regions which could be masked by other anthropogenic noises. The signal flow for multiband processing is shown in Fig. 5. An input time signal $x(n)$ is fed to a filter bank where there are M finite impulse response (FIR) filters $H_M(z)$ of order 16. The first filter $H_1(z)$ is a low-pass filter, the last one $H_M(z)$ is a high-pass filter and the ones in between $(H_2(z), \ldots, H_{M-1}(z))$ are band-pass filters. Coefficients are calculated by weighting the ideal filter response with a hamming window. Each output signal is then shifted to the baseband by multiplying the filter output with a carrier whose frequency is equal to multiples of the bandwidth $\Omega = \frac{2\pi f_{BW}}{f_s}$ of each filter, where $f_{BW} = f_s/(2\cdot M)$. The first signal, as depicted in Fig. 5, will directly pass through since it is already in the desired base-band. Each output signal is then decimated to $L = f_s/M$ so to have

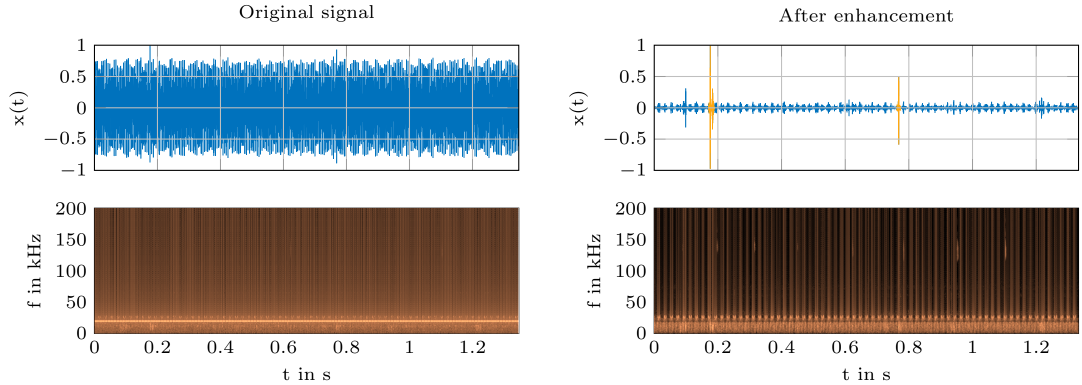

Figure 4: Click discovery after applying the transient enhancement algorithm on a recording. Tonal energies in the signal are reduced so that transients dominate. Possible click detections are depicted in yellow on the upper right plot.

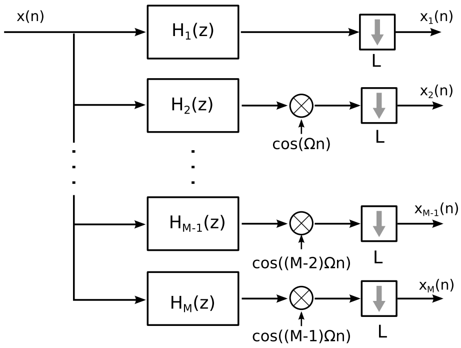

Figure 5: Signal flow of the filter bank. It can be noted that the carrier frequencies are multiples of the resulting bandwdidth of the signal.

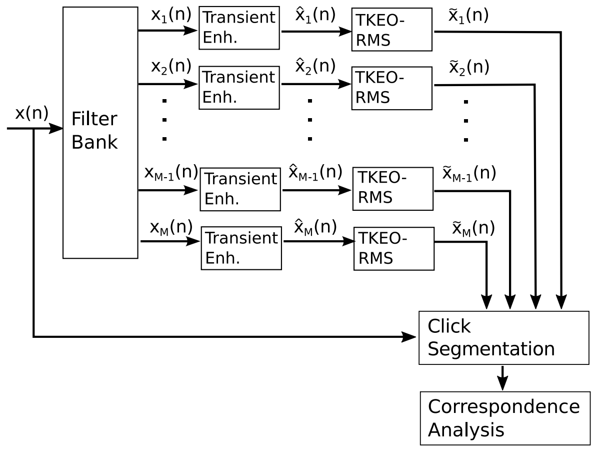

Figure 6: Block diagram of the processing algorithm for noise removal and click detection based on frequency bands. Transient enhancement is applied on each band before detection and clicks are segmented from the original signal.

a single sample rate. Fig. 7 shows an example of click detection and transient enhancement after application of the filter bank with $M = 5$ to the original signal of Fig. 4. It can be shown that the clicks in the higher frequency parts are now differentiable and clearly visible. After enhancing the click regions, clicks can be correctly identified in all of the frequency bands as in Fig.7.

Click Grouping for Different Bands

Current state-of-the-art methods for detection and classification of cetacean clicks take into consideration the mean frequency or the center frequency of each click. However, the frequency response of the clicks of individuals is not stationary in the sense that a sole mean can characterize their frequency response effectively [9]. Another used feature is the spectral spread which describes the spread energy along active frequencies of the click. However, some clicks have different frequency modulated parts, thus making it unfeasible for a generalized detection.

The correspondence analysis of Algorithm 1 compares a set of clicks in a successive manner in order to group them according to correlating characteristics [2]. Once click positions are identified, the location of the clicks, their energy level, the cross-correlation peak value are compared so that the click segments could be grouped according to its respective click train. For this study, each successive click in an incoming train is compared taking in consideration that its frequency contents remain in the same band for successive active clicks. Fig. 8 shows the output of the algorithm, where two active click trains (green in the low frequencies and magenta in the high frequencies) are detected in a recording where a Harbour Porpoise appears to be sighted. Blue detections seem to be uncorrelated to the others, and may represent echoes.

Conclusions

This paper specifies technical fundamentals for a real-time implementation of a system to detect, segment, characterize and group clicks of different cetacean species into click trains which could represent an active individual of a particular species. The current results for detection and extraction of clicks of selected species represent a first approach to determine how individual characteristics of clicks differ between species. A filter-bank based time domain click detection analysis algorithm has been implemented to account for different spectral regions and to mitigate roll-off influence in the energy content of the signals. A percussive separation algorithm has been ap-

Data: Segments of clicks
Result: Differentiated click trains
Initialize group of clicks Current click is **alpha** click;
while *Still clicks to analyze* **do**
  Get alpha click features;
  Get next click features;
  if *features correspond* **then**
    Click is part of the group;
    if *alpha click too far away* **then**
      Currrent click becomes **alpha** click;
    else
    end
  else
    Next click does not correspond, leave for next iteration;
  end
end

Algorithm 1: Click correspondence for active clicking individuals [2]

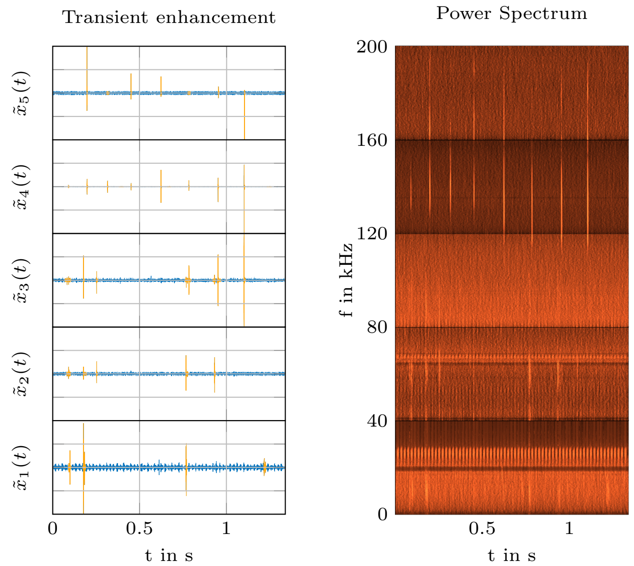

Figure 7: Extracted signal and clicks for the different frequency bands after enhancing the transient parts. Clicks are detectable in both frequency and time respresentations. Detected clicks in the time domain are colored yellow.

plied to enhance transients so to obtains a better time domain resolution for the click detector.

Further work

The frequency bands for detection can be fixed to particular frequencies allowing for a better characterization of the click sources, in which particular bands can be working in parallel to extract clicks for each species. A compound work introducing clicks from other marine mammals which exhibit the same transient-like characteristics can be modeled according to the fundamentals presented in this project. Source separation methods like the one mentioned in this work could discriminate between sweeps and clicks, thus providing event information.

Although some clicks from different species do overlap in the frequency domain, a simple discriminating function, for example based on the spectral distance between an incoming click and the mean spectra will classify the click to a particular species, and can help in the determination

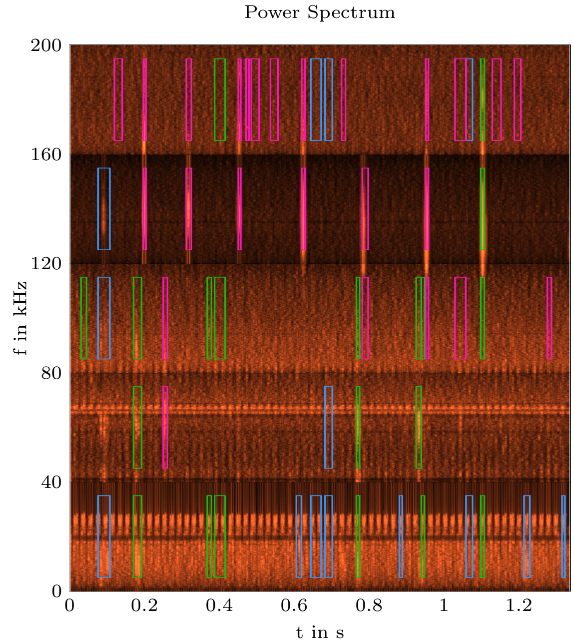

Figure 8: Extracted clicks and groups for an example signal. Clicks are differentiated in two groups, the green rectangles correspond to the low frequency clicks of the Harbour Porpoise and the magenta the higher frequency clicks. Blue rectangles show clicks which where detected, but no correspondence between the click groups was found.

of differences in individuals.

References

[1] J. H. Brown, “On the relationship between abundance and distribution of species,” *American naturalist*, pp. 255–279, 1984.

[2] C. de Obaldıa Pastor, G. Simkus, and U. Zölzer, “Estimating the number of sperm whale (physeter macrocephalus) individuals based on grouping of corresponding clicks,” in *41. Jahrestagung für Akustik (DAGA 2015), Nürnberg*, Deutsche Gesellschaft für Akustik (DEGA), 2015.

[3] C. Oliveira, M. Wahlberg, M. A. Silva, M. Johnson, R. Antunes, D. M. Wisniewska, A. Fais, J. Gonçalves, and P. T. Madsen, “Sperm whale codas may encode individuality as well as clan identity,” *The Journal of the Acoustical Society of America*, vol. 139, no. 5, pp. 2860–2869, 2016.

[4] D. P. Nowacek, L. H. Thorne, D. W. Johnston, and P. L. Tyack, “Responses of cetaceans to anthropogenic noise,” *Mammal Review*, vol. 37, no. 2, pp. 81–115, 2007.

[5] J. M. Hovem, S. Yan, X. Bao, and H. Dong, “Modeling underwater communication links,” in *Sensor Technologies and Applications, 2008. SENSORCOMM’08. Second International Conference on*, pp. 679–686, IEEE, 2008.

[6] E. Pirotta, R. Milor, N. Quick, D. Moretti, N. Di Marzio, P. Tyack, I. Boyd, and G. Hastie, “Vessel noise affects beaked whale behavior: results of a dedicated acoustic response study,” *PLoS One*, vol. 7, no. 8, p. e42535, 2012.

[7] D. Fitzgerald, “Harmonic/percussive separation using median filtering,” in *Proc. of the 13th Int. Conference on Digital Audio Effects (DAFx-10), Graz, Austria*, Dublin Institute of Technology, 2010.

[8] U. Zölzer, *DAFX: Digital Audio Effects*. John Wiley & Sons, 2011.

[9] D. Johnston, M. McDonald, J. Polovina, R. Domokos, S. Wiggins, and J. Hildebrand, “Temporal patterns in the acoustic signals of beaked whales at cross seamount,” *Biology Letters*, vol. 4, no. 2, pp. 208–211, 2008.