Sperm Whale Detection Process

Wilfried Beslin, October 2019

# Overview

The algorithm for the sperm whale detector (SWD) essentially follows the same structure as the beaked whale detection code, but some steps are implemented very differently for more efficient performance with sperm whale clicks. The basic steps consist of the following:

1. Do initial click detection
2. Compute click parameters and apply discrimination criteria to isolate candidate sperm whale clicks
3. Detect sperm whale events based on number and density of clicks
4. Manually validate sperm whale events

Steps 1-3 are run automatically using the “SWD\_master” script. As with BWD, step 4 is a separate interactive step run with “SWDValidate\_master”.

# Click Detection

Unlike BWD, initial click detection does not depend on Triton – this step is built in to the main code. However, the click detection process itself uses unaltered source code from another application (CABLE, documented in Beslin et al., 2018). This means that the audio import process at this stage is handled in a rather particular way.

## Audio Pre-processing for Click Detection

The steps for processing WAV files for click detection are outlined below. Note that this is only to obtain the start and end times of clicks; waveforms for later discrimination are not extracted at this stage.

* Divide audio files into segments
  + For duty-cycled data, individual recordings may each be a segment
* Remove quiet periods at start and end of each segment
  + That is, periods at the beginning or end where the waveform x is a constant 0, assuming no DC offset
  + This is to overcome situations where, for instance, duty cycles may have started to record prematurely or ended too late, i.e. while the AMAR was “off” and not sensitive to sound. These periods can contaminate noise level estimation during click detection.
* Resample each segment to 48 kHz
  + This is done mainly to improve performance, and is possible because most energy in sperm whale clicks is below 24 kHz. Has the added benefit of reducing likelihood that other odontocete clicks are detected.
  + An FIR antialiasing filter is applied before resampling
    - Order = , where is the reduced resampling ratio (e.g. for downsampling 96 kHz to 48 kHz)
* Filter noise from resampled segments
  + 10th order highpass Butterworth filter
  + 2 kHz half-power frequency cutoff

## Click Detection using CABLE

Start and end times of clicks are detected using the same code from CABLE (Beslin et al., 2018). This algorithm is very similar to the one used by PAMGUARD, but it has a few extra features to more accurately capture multi-pulsed sperm whale clicks. It is documented in full in Appendix A, section B of Beslin et al. (2018), but basically follows these steps:

* The algorithm examines each sample in the timeseries in sequence, and begins in an ’off’ state (i.e. it starts assuming no detection)
* For each sample, a recursive instantaneous estimate of signal and noise power is computed. In other words, signal and noise power for one sample are estimated based on the power estimates of the previous samples.
* If the signal-to-noise ratio exceeds a certain threshold for a given sample, then the detector switches to the ‘on’ state where it marks the sample as part of a click detection
* The ‘on’ state is maintained for subsequent samples until the signal-to-noise ratio falls below a second threshold, after which it returns to the ‘off’ state
* After this process has been completed for all samples, the periods where the detector switched states (which correspond to estimated start and end times of clicks) are revised to remove spurious detections (i.e. very short clicks) and to ensure that clicks have not been broken up by minor fluctuations in power.

SWD uses the same default click detection parameters as in CABLE except for the minimum click duration (which is effectively 50 µs this case, versus 2 ms in CABLE). The default “on” and “off” SNR thresholds are also different (20 versus 10 for “on”, 2 versus 1 for “off”)

# Click Discrimination

Using the times determined from the click detection step, click waveforms are isolated from the raw audio file and several features are computed for later discrimination. This part is very similar to BWD: first, waveforms are extracted including a 1 and 4 ms buffer before and after the click, respectively. For each click, a noise sample is also taken as the first 5 ms before the start of the click. Both click and noise samples are filtered using a 2-120 kHz bandpass Butterworth filter. Spectra are also computed using an FFT window size of 512 points by default, though users can change this. Following this, each feature is evaluated to ensure it is within the range expected for sperm whale clicks.

The list of features and the thresholds applied to each click is dictated by the file in the following path within the SWD directory:

*DetectionCriteria\ClickDiscrimParams.xlsx*

At present, the list of features and thresholds used is as listed in **Table 1**. These were chosen mainly by examining the feature distributions of 243 sperm whale clicks from various AMAR recordings along the Scotian shelf. Further performance analysis was conducted at the event detection stage (see below).

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1**: list of criteria used for sperm whale click discrimination | | | |
| **Criterion** | **Description** | **Threshold (Min)** | **Threshold (Max)** |
| Fpeak | Peak frequency (kHz) | 0 | 20 |
| F0 | Centroid frequency (kHz) | 0 | 20 |
| bw10db | Magnitude of 10 dB bandwidth (kHz) | 0 | 25 |
| bw10dbLower | Lower frequency of 10 dB bandwidth (kHz) | 0 | 15 |
| bw10dbUpper | Upper frequency of 10 dB bandwidth (kHz) | 0 | 28 |

# Event Detection

From the candidate sperm whale clicks, each audio segment is evaluated to determine if a sperm whale presence event has occurred or not. This is based on criteria that evaluate the number and separation of clicks. At present, the list of features and thresholds used is as listed in **Table 2**.

|  |  |  |
| --- | --- | --- |
| **Table 2**: list of criteria used for sperm whale event detection | | |
| **Criterion** | **Description** | **Threshold** |
| MaxMedianICI | Maximum value of the median of the inter-click interval, in seconds | 2 |

Event detection performance was tested by running the code on 247 AMAR recordings from various sites along the Scotian Shelf, each about 1 minute in duration. These recordings were manually validated for sperm whale presence; about half had confirmed sperm whale presence, and the other half had effective sperm whale absence.

# Manual Event Validation

This step works very much the same as in BWD, with a few small differences. Each event is presented one at a time with a series of figures which aid the user in confirming whether sperm whales are present or not. Figures consist of the following:

* Plots summarizing the peak frequency, ICI, and mean spectrum of clicks within that event
* The waveform and spectrogram of an individual click, along with a plot showing the occurrence and intensity of all clicks within ±2 ms of the target click

Users may cycle through each candidate sperm whale click in the event. Once a decision on presence/absence is made, the next event is loaded. Unlike with BWD, only one species is targeted in this case, so there is no option to discriminate clicks dynamically, and reference spectra are not displayed.

As with BWD, segments that had positive detections during automatic event detection may be merged into larger events based on time gaps or calendar durations.

# References

Beslin, W.A.M., Whitehead, H., and Gero, S. (**2018**). “Automatic acoustic estimation of sperm whale size distributions achieved through machine recognition of on-axis clicks,” J. Acoust. Soc. Am. **144**, 3485-3495. [doi: 10.1121/1.5082291](https://doi.org/10.1121/1.5082291)