# Spectral peak frequencies measurement sets for dolphin whistles

**Name and contact**: Pina Gruden, Institute of Sound and Vibration Research (ISVR), University of Southampton; pg3g12@soton.ac.uk

The purpose of the provided measurement set and the corresponding interpolated ground truth data of dolphin whistle contours, is to provide a common measurement set on which the automated algorithms for whistle contour extraction can be tested.

This dataset was used in the following publications:

Gruden, P. & White, P. 2016. Automated tracking of dolphin whistles using Gaussian mixture probability hypothesis density filters. *Journal of the Acoustical Society of America,* 140, 1981 - 1991.

Gruden, P. & White, P. Submitted. Automated extraction of dolphin whistles – a Sequential Monte Carlo Probability Hypothesis Density (SMC-PHD) approach.

Below is a description of how the spectral peak frequency measurements and interpolated ground truth data were obtained.

## Spectral peak frequency measurements

This measurement set contains spectral peak frequency measurements obtained from the files released for the 5th Workshop on Detection, Classification, Localization and Density estimation of marine mammals (DCLDE) 2011, found on the MobySound archive page (http://www.mobysound.org). Dataset contains six odontocete species; long-beaked common dolphin (*Delphinus capensis*), short-beaked common dolphin (*Delphinus delphis*), melon-headed whale (*Peponocephala electra*), spinner dolphin (*Stenella longirostris*), Atlantic spotted dolphin (*Stenella frontalis*) and bottlenose dolphin (*Tursiops truncatus*), and specific files used are listed in a document titled *"File_list.xlsx"*. The files that were selected for the measurements have corresponding analyst annotated data files (.bin) that can be found in the MobySound 5th Workshop 2011 folder, which also contains additional information about the original dataset.

## Procedure for obtaining the measurements

All raw data was first re-sampled to 192 kHz. After re-sampling, pre-processing was applied to the data in order to reduce the background noise and interfering signals. A pre-processing scheme was adapted from Gillespie *et al*. (2013) and was applied with a sliding window that was 2048 points long and had 50% overlap, resulting in 93.8 Hz spacing between frequency bins. Within each window the following steps were performed as described in Gillespie *et al*.

(2013): first echolocation clicks were removed by applying a weighting function; then spectrogram was computed on a decibel scale, using 2048 point Hann window, and spectral peaks were enhanced by applying normalization across frequency based on a 61 point median filter; after that the normalization across time using exponential moving average was performed in order to remove persistent tones from the spectrogram.

In each window, after the noise was removed, spectral peaks were determined by identifying all frequencies whose normalized magnitude exceeded 8 dB threshold. Only frequency bins between 2 and 50 kHz were searched for peaks, since most dolphin whistles will lie within this range.

A sample code used to identify spectral peaks can be found in the "*Sample_code*" folder. Measurement sets obtained as described above can be found in the "*Measurement_sets_MATfiles*" folder, where they are sorted in subfolders by species. Measurement sets are named with a following naming convention: "*Zset_audiofilename.mat*". Each measurement set is a .mat file and it contains the following:

- file – name of the .wav file that measurements were obtained from
- fs – sampling rate
- win_width – window length in samples
- slide_incr – time increment- increment between consecutive windows in samples (*e.g.*, for window length 2048 samples and 50% overlap between consecutive windows, the time increment would be 1024 samples. )
- Zset – a cell array containing spectral peak measurements; each cell corresponds to a time step and contains spectral peak frequency measurements in Hz for that time step. Time step is defined as *slide_incr/fs* (For example, in a cell array with 10 cells, *slide_incr* of 1024 and *fs* of 192 kHz, the time steps would be 0.0053, 0.0107, 0.0160, 0.0213, 0.0267, 0.0320, 0.0373, 0.0427, 0.0480, 0.0533 s). The spectral peaks measurements are a mixture of whistle contour measurements and clutter (measurements not corresponding to whistles).

## Ground truth data

In addition to the spectral peak measurements for each file, we also include interpolated ground truth data obtained from the analyst-annotated files in the DCLDE dataset. The ground truth files have been processed so that the time increments and corresponding frequency measurements match the parameters used above. To achieve this the ground truth data is interpolated using linear interpolation. In addition, each ground truth whistle is evaluated in terms of its duration and SNR. Each ground truth whistle was only expected to be detected if its duration exceeded 150 ms and if its SNR exceeded 10 dB above the background noise for at least one third of its duration (following Roch *et al*. (2011)). The background noise was determined across 5 s windows with a median filter. Each ground truth whistle meeting these criteria was termed valid.

The interpolated and evaluated ground truth data can be found in the respective

.mat files in the folder called "*Interpolated_Ground_Truth*". The ground truth data is sorted in the subfolders by species and the naming corresponds to the original audio file name, with a following naming convention: "*GT_audiofilename.mat*". Each .mat file contains the following:

- GT – a structure with three fields, where each whistle is given in a separate row; the fields are: *time* (a vector of times in s from the beginning of the file), *freq* (a vector of whistle's frequencies in Hz), *valid* (indication whether the ground truth whistle met the selection criteria described above (1) or not (0)).
- Nvalid – number of valid whistles (whistles that met the selection criteria) in a file.

**Comparison of detection algorithms that operate on spectral peaks**

As mentioned above, it is our hope that this dataset can facilitate the comparison of detection algorithms that operate on tracking from spectral peaks. To this end we are also including performance results for detected whistles that were longer than 150 ms (file "*GMPHDandSMCPHD_performance.xlsx*" in the "*Detectors_performance*" folder). The results are given for two multi-target tracking algorithms, based on the Probability Hypothesis Density (PHD) filter, that were used to extract whistle contours from the given measurement dataset. See Gruden and White (2016) and Gruden and White (submitted) for details on the algorithms. The performance was evaluated in terms of precision (percentage of detections that are correct), recall (percentage of the expected valid whistles that are retrieved), coverage (average percentage of the ground truth whistle that is covered by detector), fragmentation (average number of detections per ground truth whistle) and average deviation from the annotated whistle path, as described in Roch *et al*. (2011). The ground truth data used to evaluate the performance can be found in the folder "*Interpolated_Ground_Truth*".

**Viewing and Importing Spectral peak frequency measurements data with an open source software**

For the ease of access, the Matlab versions of the measurement sets are converted to .TXT files, that follow the same naming convention as described above and can be found in the folder "*Measurement_sets_TXTfiles*". A sample code is provided that shows how these conversions were achieved and a code that allows one to import these .txt files back into the Matlab. This code can be found in the "*Sample_code*" folder, in a subfolder called "*ConvertToandRead_TXTfiles*".

Additionally, a python functions are supplied, that read and upload the datasets into the Python open source program. These can be found in the "*Sample_code*" folder, in a subfolder called "*ConvertTo_Python*".

## References

GILLESPIE, D., CAILLAT, M., GORDON, J. & WHITE, P. 2013. Automatic detection and classification of odontocete whistles. *Journal of the Acoustical Society of America,* 134**,** 2427-2437.

GRUDEN, P. & WHITE, P. 2016. Automated tracking of dolphin whistles using Gaussian mixture probability hypothesis density filters. *Journal of the Acoustical Society of America,* 140, 1981 - 1991.

GRUDEN, P. & WHITE, P. Submitted. Automated extraction of dolphin whistles – a Sequential Monte Carlo Probability Hypothesis Density (SMC-PHD) approach.

ROCH, M. A., BRANDES, T. S., PATEL, B., BARKLEY, Y., BAUMANN-PICKERING, S. & SOLDEVILLA, M. S. 2011. Automated extraction of odontocete whistle contours. *Journal of the Acoustical Society of America,* 130**,** 2212-2223.