Robust Automatic Detector And Feature Extractor For Dolphin Whistles

Yoel Bud[1,+], Guy Shkury[1,+], Roee Diamant[2], Yotam Zuriel[3], and Aviad Scheinin[3]

[1] Andrew and Erna Viterbi Faculty of Electrical Engineering. Technion Israel Institute of Technology
[2]Hatter Department of Marine Technologies, University of Haifa, Israel
[3]Morris Kahn Marine Research Station, University of Haifa, Israel
[+]Equal contribution

***Abstract*— A key in Dolphin's conservation efforts is population estimation in their natural environment. A common method for mapping Dolphin's appearance is the detection of their vocalizations. In this paper, we propose a novel detection technique for Dolphin's whistles, referred to as ECV (Entropy, Correlation, and Viterbi algorithm). ECV is a robust detector of low complexity that automatically detects dolphin's whistles and extracts their spectral features, using a single receiver with only a few system parameters. The method employs a chain of decisions based on spectral entropy and time-domain correlation followed by constrained Viterbi algorithm to extract the whistles' features. Simulation results as well as performance over real recordings shows a good trade off between detection and false alarm, that compares well with the widely used PAMguard system.**

I. INTRODUCTION

Conservation of top predators in an ecosystem is crucial for the existence of the entire food chain. Among the most interesting marine top predators are dolphins, a species whose diversity reflects on the health of the entire marine environment [1, 2]. Off Israel's coast, invasive species and human interaction (e.g. fishing activities and construction and operating gas rigs) pose threats on the strive of dolphins, and raise concerns among conservationists.

A common way to estimate the impact of such threats on dolphins' populations is to find indication of their existence or non-existence across large surveyed areas, mostly by detecting Dolphin's localizations. To that end, man-in-the-loop methods (e.g., [3]) may be inefficient, and the challenge is to develop automatic detection with high precision. In particular, such a tool will be useful to map the geographical distribution of dolphins, their daily routine, etc. In this work, we focus on the automatic detection of Dolphins' whistles. The common methods for detection of dolphins' whistles include software solutions like PAMGuard [4], which performs detection by frequency domain amplitude, and methods based on the analysis of time-frequency spectrum images [5]. However, these methods require expert supervision to manually adapt detection thresholds, and may not fit the case of long term data analysis.

Tracking the population of dolphins using their acoustic emissions requires to overcome three main challenges: detection within strong ambient and man-made noise, feature extraction, and classification. Erbs et al [6] used an array of towed hydrophones in order to record four types of dolphins, and PAMguard in order to detect and classify their whistles. PAMguard whistle detection is highly configurable, it mostly relies on the following modules, Energy sum detection , Spectrogram correlation and a Matched filter detector configured to the types of whistles expected. Mahdi Esfahanian et al [7] compared two methods to classify and detect different types of whistles produced by Bottlenose dolphins. The first relies on Fourier Descriptors and second on temporal and spectral features of the whistles. Features of whistle spectral contour lines proved to be effective for whistle classification. In their paper, suction-cup hydrophones were used to record underwater acoustics such that high SNR values are obtained. The features were analyzed using support vector machine (SVM) and K-nearest neighbors (KNN) classifiers. Oswald et al [3] classified nine different species using a similar methodology of manually detecting whistles and extracting their spectral information. They found that the most effective spectral parameters are: minimum frequency, maximum frequency, start frequency, end frequency, frequency range, and time duration.

Automation of the feature extraction process is considered to be a challenging task because of the channel's high and time-varying ambient noise, and because of the movement of the dolphins themselves that distorts the received signals. Confronting with these challenges, Song et al [8] used an array of hydrophones, and created an automated system to detect Yangtze Finless Porpoise using Hilbert transformation as a feature extraction method. Kohlsdorf et al [9] proposed a probabilistic method of tracing distorted spectrogram contours. This method is based on a thorough investigation of fundamental units in these signals. While these methods achieve good detection rates, their results are confined to specific scenarios of high signal-to-noise ratios and do not handle practical challenges such as man-made acoustic noises, and channel's time variations.

In this paper, we propose an automatic detector for Dolphins' whistles, referred to as Entropy, Correlation, and Viterbi detector (*ECV*). ECV is aimed to detect all kinds of Dolphin whistles, and does not require training data.

The method does not consider a model for the whistle, and other than assuming a bound on the rate of frequency change, self-adapts to the statistics of the ambient noise. Furthermore, ECV is not just a detector, but also provides an estimate for the spectral features of the detected whistle. As such, our contribution is twofold:

1) A robust detector of low complexity to automatically detect dolphin whistles.
2) A novel way to extract spectral features from dolphin whistles.

II. Model and Assumptions

Our system model includes a single hydrophone recording opportunistic acoustic emissions. The recordings are expected to include large noise portions and some dolphin whistles. To cover large areas, the surveying vessel is moving while collecting measurements. Consequently, the deployed hydrophone is dragged behind the vessel, and the recording includes noises from both the boat's motor and flow noises. The setup is illustrated in Fig. 1.

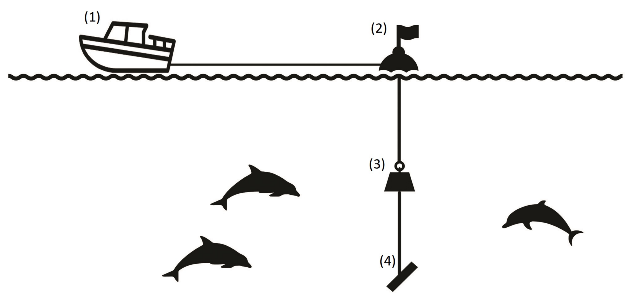


Fig. 1: An illustration of the recording technique used to record dolphin whistles. The surveying vessel (1) moves in order to cover large geographical area. The hydrophone (4) is connected to the vessel using a cable which is attached in two connection points: to a buoy (2) and a weight (3). The buoy keeps the cable floating behind the vessel, and the weight is attached close to the hydrophone.

Our noise model includes ambient isotropic noise and non-isotropic noise terms. The former is modeled as an i.i.d Gaussian process, while the latter is modeled by wideband impulsive-like transients as well as correlated low-frequency noises. To formalize, for an input signal y(t) containing Dolphin's whistle, we have:

$$y(t) = d(t) + n(t), \tag{1}$$

$$n(t) = n_g(t) + n_h(t) + n_u(t) \tag{2}$$

where $d(t)$ is a dolphin whistle, $n_g(t)$ and $n_u(t)$ represent i.i.d Gaussian noise and a noise transient, respectively, and $n_u(t)$ represents low frequency artificial noise.

Dolphins produce several types of acoustic signals. The most commonly observed are *Whistles* and *Clicks* [10]. The former is characterized by a long emission with time-varying chirp-like spectral content and is said to be used for communications [11], while the latter is a wideband short signal emitted in almost constant periods and is said to be used for ranging and for forging [12]. In this work, we focus on the identification of dolphin's whistles. We assume dolphin whistles are within a limited bandwidth of 5-24 KHz, are continuous signals, and their duration ranges between 200 milliseconds to 2 seconds [13].

III. The ECV Detector

A. Key Idea

The ECV algorithm composes 3 main stages (see illustration in Fig. 2), namely an entropy detector, a correlation detector, and a constrained Viterbi algorithm— also used for feature extraction. All stages are built as a detection mechanism, aimed to reduce false negative decisions. This structure is chosen by our intuition that Dolphin noises are stationary signals, in contrary to the noise. Hence, the system starts with a band pass filter aimed to increase time-domain signal-to-noise ratio and to remove possible correlated low frequency noise components that may effect the entropy and correlation detectors.
After this preprocessing stage, comes a spectral entropy detector followed by a temporal correlation detector. Our entropy detector uses the continuity of dolphin whistles in the frequency domain in order to detect a decrease in the spectral entropy. The time correlator, which utilizes the continuousness of the whistle, is then applied. The result is a time segment, suspected to include a Dolphin's whistle.
Finally, detection verification and feature extraction are performed using a constrained Viterbi algorithm. In particular, we feed in the spectrum of the detected time segment, and regard the time samples as observations and the frequency bins as states. Then, the emission belief is heuristically set as the signal's normalized spectrum, while the transition probability is set to allow a maximum value for state/frequency transitions. The result of this constrained Viterbi algorithm is a track that follows the spectral contour of the whistle, which can also be used to extract the spectral features for classification. In the following chapters, we describe in details the structure of ECV.

B. Entropy detector

The instantaneous spectral entropy of a time-frequency power spectrogram S(t,f) is:

$$P(t,m) = \frac{S(t,m)}{\sum_f S(t,f)} \tag{3}$$

And the spectral entropy at time t is:

$$H(t) = \sum_{m=1}^{N} P(t,m) \log_2 P(t,m) \tag{4}$$

Observing (4), we note that $H(t)$ is high for a random signal like n(t) in (2), but is low for a stationary signal, as we expect a Dolphin whistle to be. Hence, a decrease in $H(t)$ may indicate the existence of a signal. To find such a decrease, in ECV we slice the recorded signal from the

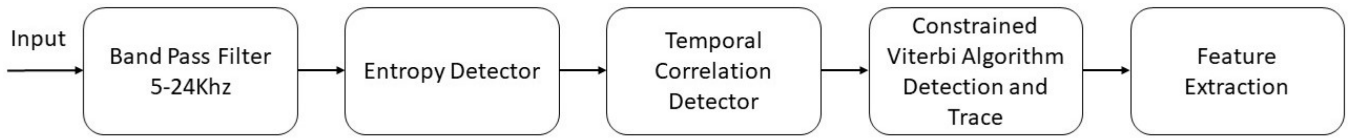


Fig. 2: A block diagram for the operation of ECV

channel using a sliding window, and compute $H(t)$ for short time frames. The sequence of $H(t)$ values is then compared to a detection threshold.

Each sequence of $H(t)$ is adaptively normalized by the mean and standard deviation values of $H(t)$ determined as 'noise', μ_H and σ_H, respectively, such that:

$$\widehat{H}(t) = \frac{H(t) - \mu_{Hnoise}}{\sigma_{Hnoise}} \tag{5}$$

Assuming i.i.d Gaussian distribution for the spectral entropy of the noise, the threshold, Th is calculated assuming normal distribution, such that [14]:

$$P_{\text{fa}} = \frac{1}{\sqrt{2\pi}} \int_{\mathrm{Th}}^{\infty} e^{\frac{t^2}{2}} dt \tag{6}$$

where P_{fa} is a user defined parameter for the desired false alarm. An $\widehat{H}$ value lower than the threshold is then tagged as a potential whistle. To validate detection, segments marked as potentially containing whistles are filtered such that a segment shorter than a minimum duration α or larger than a maximum β, is discarded.And α and β are set by the expected whistle duration [13].

C. Temporal correlation detector

Temporal correlation detection is done by correlating adjacent segments of the recorded buffer flagged by the spectral entropy detector. Let $\vec{x}(t), 0 < t < T$ be a time domain buffer of duration T sec, and let ω represent the correlation window size. Then, the correlation parameter is determined as

$$C(t, \omega) = \text{Max} \frac{\text{Corr}(\vec{x}(t), \vec{x}\,(t + \omega))}{\int_t^{t+\omega} |\vec{x}(\tau)|^2 d\tau} \tag{7}$$

Assuming the change in frequency over time is subtle for a dolphin's whistle but is significant for an i.i.d noise, $C(t, \omega)$ is expected to increase if $\vec{x}(t)$ includes a whistle. For fine resolution, the temporal correlation (7) is performed for a sliding window $\vec{x}(t)$, where ω is a user defined parameter that trades off resilience to noise components (increasing with ω) with sensitivity to signal variations (decreases as ω increases).

While the output of the correlator in (7) identifies well a target, it is sensitive to strong bursts of energy. Thus, in ECV, we operate the correlator after the entropy detector. Another reason for this is that the correlation method is better than the entropy detector for identifying the starting and ending points of the whistles. Detection using the correlation output involves a predetermined threshold δ, where δ is chosen from the receiver operating characteristics (ROC) in Fig.4b.

D. Constrained Viterbi Algorithm

We find that the sequence of entropy and temporal correlators produce good detection results. However, it may also induce significant false negatives for correlated noise components, produced by signals e.g., boats' thrusters. For this reason, we validate detection by testing if the detected signal fits a dolphin's whistle. That is, we test that the signal is not constant in time and in frequency. To that end, we use a constrained Viterbi algorithm.

The Viterbi algorithm is traditionally used for probability analysis in long observation sequences, and is mostly used in communication applications [15]. Here, we employ it to track over spectral lines. This is performed by considering the frequency bins as states in a Hidden Markov chain, and the time samples as observations. For the emission matrix we use the normalized time-frequency spectrum matrix $\vec{x}(t, f)$ containing the time window suspected to include a whistle. The output of the Viterbi algorithm is a probability to find a continuous path vector, $\vec{P}(t)$, that represents frequency bins over time assumed to belong to a whistle. Since the Viterbi algorithm is geared to find the most probable state path [15], running it on the spectral matrix $\vec{x}(t, f)$ yields a spectral contour line with the largest continuous spectral energy.

Since a dolphin's whistle is expected to be continuous in frequency, we do not expect *jumps* in the frequency domain. That is, the spectral line that corresponds to the time-frequency characteristics of the whistle should not include large variations within frequency binsbetween consecutive time instances. To force such a solution, we constrain the Viterbi algorithm by difining the state transition matrix to be

$$T_{ij} = \begin{cases} 1/\kappa & \text{for } i - \kappa/2 < j < i + \kappa/2 \\ 0 & \text{o.w} \end{cases} \tag{8}$$

where i and j are two frequency bin states, and κ is the maximum number of states or frequency bins allowed for a transition between two consecutive observations. We set κ by the user's expectation of the rate of change in the spectral content of the dolphin's whistle. Note that the Viterbi algorithm will always issue a path $p(t)$. Hence, to removefalse detections we compare the path probability, i.e., the average of $p(t)$, to a predetermined confidence value,γ.

E. Feature extraction

Since for dolphin's census, not just the detection of dolphins is of interest but also their classification, in ECV we offer a way to extract the whistle features. While some classifiers, e.g. convolutional neural networks [16], can operate directly on the raw signal, still feature extraction is likely to improve classification performance. Once the trace $p(t)$ is obtained, the spectral features can be easily extracted from the evaluated contour line. Specifically, we evaluate the maximum and minimum frequencies of the whistle, the start and end frequency, and the signal's duration.

F. Example of Operation

An example of the operation of ECV is shown in Fig.3a and Fig.3b. The figures illustrate the process of detection and tracing of two dolphin whistles. Fig.3a shows the output of the energy and correlation detections. We observe that the entropy detector is more robust while the correlation detector is more sensitive to noise. In turn, the entropy detector smooths the signal and hence have low resolution, while the correlation detector have fine time resolution. Considering these differences, we use the entropy detector to detect the whistles, and the correlation detector to find the starting and ending point of a whistle.

Fig.3b shows the output of the constrained Viterbi algorithm for the whistle in Fig.3a. The figure shows the spectrogram that is used as an. The figure shows the spectrogram that is used as an input for the Viterbi algorithm, and its resulting estimated path of the whistle. Fig.3b also shows the probability of the estimated path. Note how high noise causes the probability to decrease, and at certain times, to diverge from the correct trace of the whistle, as seen in the left panel.

IV. Performance Evaluation and Discussion

In this section, we explore the performance of ECV. Our performance analysis is done in terms of the receiver operating characteristic (ROC) to trade off detection and false alarm rates and in terms of the accuracy of the evaluated whistle's features. To obtain enough statistics, we calculate the ROC using a simulated database, while recordings of real Dolphin whistles are used for verification.

A. Simulation Structure

To simulate the case of dolphin's whistle, we consider a large recording of 3 hours long, containing both simulated noise and simulated whistle. The noise is an i.i.d white Gaussian, while the whistle is simulated by a chirp signal. A total of 1000 chirps are placed uniformly randomly in different locations across the recorded buffer. The duration of the chirp is uniformly generated between 0.2 s and 1.5 s, and its start and ending frequencies are uniformly generated between 5 kHz and 24 kHz. This setting allows us to explore the performance of ECV for a variety of whistle-like signals. The SNR is defined by:

$$SNR = 10 \log \frac{P_{\mathrm{sig}}}{P_{\mathrm{noise}}} \tag{9}$$

where P_{sig} stands for the sum of squared values across the spectral line of the signal, and P_{noise} reflects the sum of squared values across the spectrum minus P_{sig}.

B. Detection Analysis

ECV parameters can be fitted for many different whistles characteristics and environments. In this work each parameter is chosen based on simulation results. From the simulations, the impact system parameters have on the ROCs'4 is dominant for $SNR = 0$, for $SNR > 0$ we can achieve zero false detection, and for $SNR < 0$ we can't detect whistle based on our model. The entropy detection threshold ρ sets the initial amount of noise to whistles ratio fed into the system 4a. The temporal correlation detector threshold β is significant for finding the beginning of the whistle. The smaller the threshold, the higher the accuracy in start time detection. The Viterbi Algorithm confidence threshold γ controls the level of certainty in the trace, favoring whistles with stronger SNR. In 4c we learn that the Higher the threshold the lower detection is as well as false alarms. The rate of change in the spectral content affects the trace done by the Viterbi Algorithm. Recall the Viterbi algorithm will always find a path, Higher are suitable for whistles with high frequency change rates.

Feature extraction accuracy evaluation over simulated whistles		
Parameter / Error	Mean	STD
Start time [sec]	0.03	0.02
End time [sec]	0.9	0.07
Start Frequency [kHz]	1	0.8
End Frequency [kHz]	6	4.5
Max Frequency [kHz]	0.2	0.8
Min Frequency [kHz]	3.2	3.1

TABLE I: Feature extraction based on Viterbi algorithm tracing for simulated whistles.

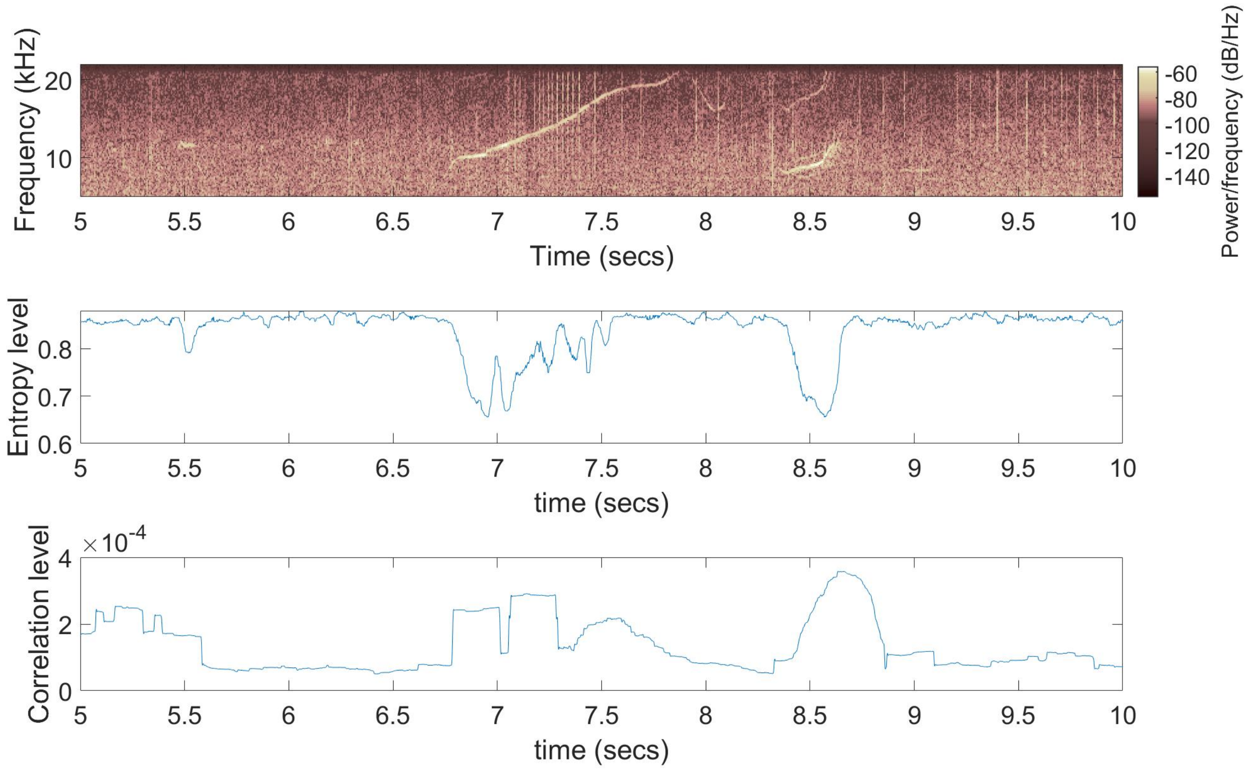
Feature extraction accuracy evaluation over real tagged whistles		
Parameter / Error	Mean	STD
Start time [sec]	0.2	0.28
End time [sec]	7	35
Start Frequency [kHz]	2.8	2.6
End Frequency [kHz]	2.7	3
Max Frequency [kHz]	1.6	2.2
Min Frequency [kHz]	3.7	3

TABLE II: Feature extraction based on Viterbi algorithm tracing for real whistles.

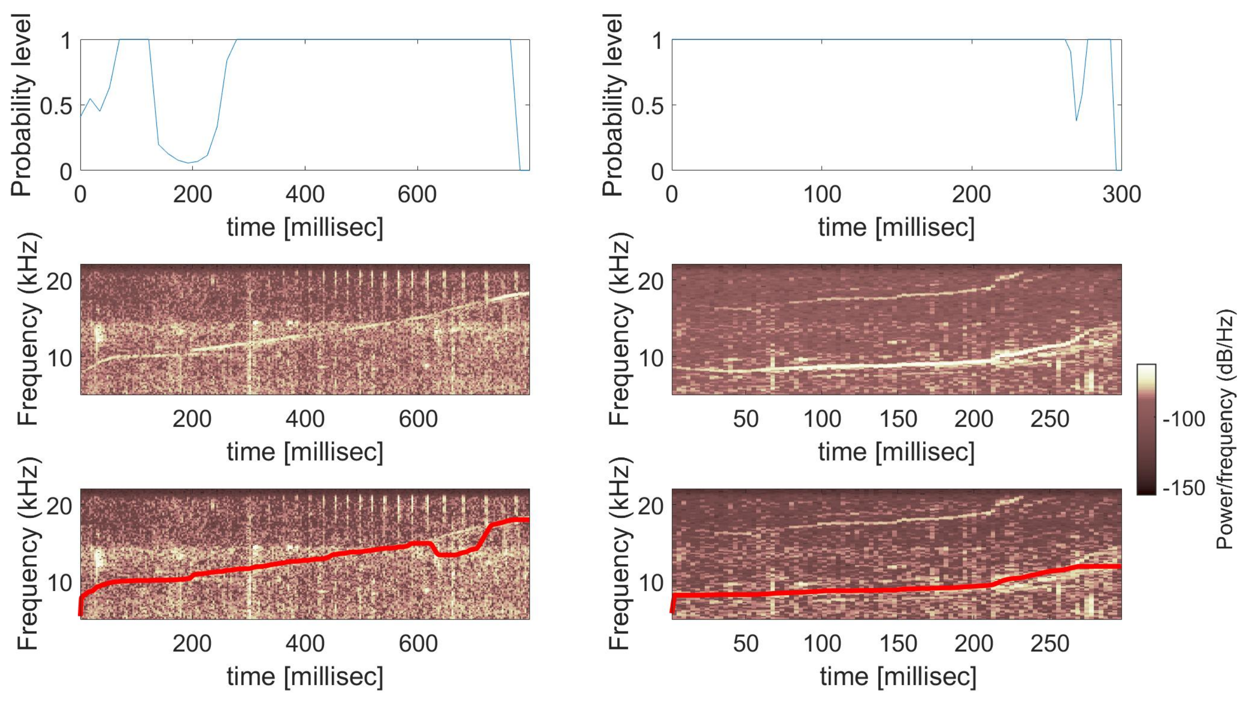
C. Analysis of Real Dolphin's Whistles

In our analysis, we have simulated a Dolphin's whistle by a chirp signal. To verify the results of the simulations, we now report results for real Dolphin's whistle. Several recording expeditions took place. The setup of these surveys is illustrated in Fig.1. We obtained a total of 5 hours of recordings all in the Mediterranean Sea across the shores of central Israel, and 4 additional hours from the Red Sea. The hydrophones used for these recordings use a sample rate of 96[KHz]. In addition, we have used one recording from the **** competition, which included recordings off the coast of France. From all these recordings, we have manually tagged and measured the features of 140 whistles.

As a benchmark, we consider a commonly used system for detection of Dolphin's whistles, namely, the Passive

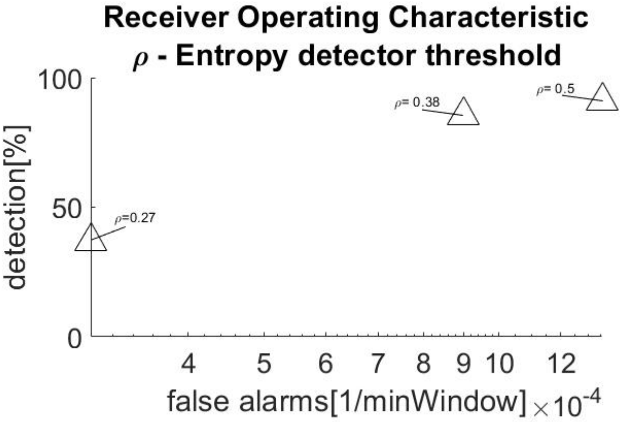


(a) The upper panel shows a spectrogram containing two dolphin whistles. The middle panel shows an entropy detector and we observe a decrease of the entropy level where a whistle exists. The lower panel shows the output of the temporal correlator, and an increase is observed where whistles are located. Recording was collected off the coast of Ashdod, Israel in May 2018.

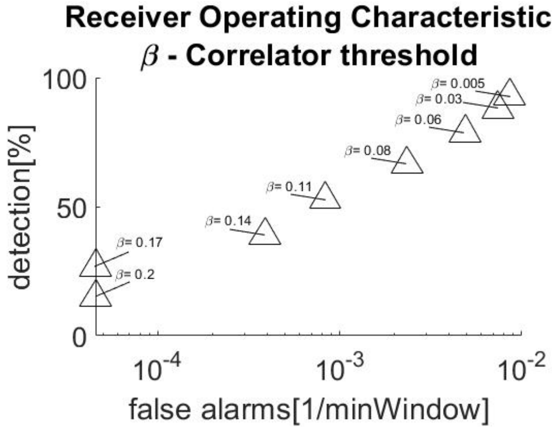


(b) Illustration of the constrained Viterbi module. The module first calculates the path probability, and then traces the whistles' contours. Notice that the algorithm will always find a path and is susceptible to deviate due to strong noise signals

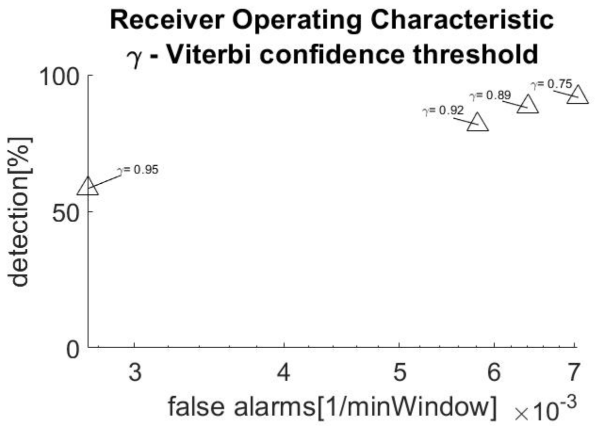
Fig. 3: This figure illustrates how ECV detects and extracts features of dolphin whistles.



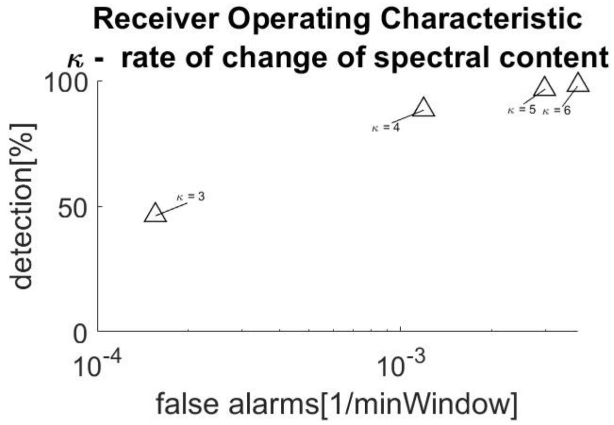
(a) ROC for the Entropy detector. For SNR=0 dB, results show a good trade off between detection and false alarm rates for threshold $\rho = 0.38$.



(b) ROC for the Correlation detector. For SNR=0 dB, good trade off between detection and false alarm rates is obtained for $\beta = 0.06$.



(c) ROC for the Viterbi Algorithm module noise threshold. For SNR=0 dB, good trade off between detection and false alarm rates is obtained for $\gamma = 0.92$.



(d) ROC for the Viterbi Algorithm detector. For SNR=0 dB, good trade off between detection and false alarm rates is obtained for $\kappa = 4$.

Fig. 4: ROC for system parameters

Acoustic Monitoring open source software (PAMGuard) [4]. PAMGuard is used extensively by marine biologists, and can be operated as an automatic detector. The system requires several set parameters. Specifically, a threshold for the signal energy, the searched bandwidth, the size of a median filter used, etc. In our analysis, we choose these parameters as the best that fitted our database. For ECV, we used the parameters obtained as the best trade off from the ROC curve for the simulated data, in particular see table III.

ECV parameters used for evaluation	
ρ	0.35
β	0.11
γ	0.85
κ	8

TABLE III: Parameters selected from ROCs' for evaluation of real whistles

The results in terms of feature extraction obtained by ECV for the real recordings are shown in Tables I,II. We observe high accuracy in terms of the start time and start frequency. Yet, the results show that ECV is not so accurate in terms of the ending time of the whistle and its ending frequency. Real dolphin whistles don't maintain a constant SNR, getting weaker at certain points over time, therefore detecting the full duration of the whistle requires the algorithm to be forgiving to sudden changes after a suspected whistle has been detected. For example, this results in high sensitivity in the drop of the entropy level but not as sensitive to the rise, as ECV can't be certain the whistle is over or just temporarily weak. This can be easily corrected running ECV twice: once as a causal system, and once when the time domain is flipped.

The results in terms of detection are described in Table IV. Upon initial parameters based only on simulations, we achieved approximately 50% detection rates but high false alarm values, PAMGUARD on the other hand was giving no false alarms but detection at rates lower than 5%, To achieve a reasonable comparison we minimally adjusted ECV and PAMGUARD parameters to find the tipping point, from there on the detection and false alarm rates diverge, detection to 0% or false alarms to high values. We observe that ECV achieves a detection rate of roughly 27%. The results also show that PAMGUARD achieves a lower detection rate of 20%. On the other hand, the results show that PAMGUARD is slightly better in terms of false alarm. This is due to PAMGuard mode of operation to correlate the received signal with some known shapes of dolphin whistles. In particular we used the configurations Sperm Whale Click and Dolphin Whistle Detection configuration. For the ROCCA classifier we used Northwest Atlantic Classifiers. These configurations are currently being used by Morris Kahn Marine Research Station, University of Haifa, Israel. Both configurations are available at the PAMGUARD website.

Evaluation over real tagged whistles		
	ECV	PAMGUARD + Rocca
True detection[%]	27	20
False detection $[1/minwindow]$	10^-2	10^-3

TABLE IV: Detection results for real Dolphin's whistles

V. Conclusions and Future Work

In this paper we introduced ECV: a novel fully automatic approach to detect and extract spectral features of dolphin whistles. ECV works by a chain of detectors starting from spectral entropy to detect stationary signals, followed by correlation detector assuming the whistle is slowly changing in frequency, and ending with a constrained Viterbi algorithm to lock onto spectral contour lines. Different than common approaches, ECV does not require man-in-the-loop intervention, and its few parameters are set by transfer learning from simulated database. ECV is therefore a robust solution to the hard problem of detecting and characterizing dolphin's whistle in a noisy sea environment.

Our simulation results showed good detection performance for SNR levels as low as 0 dB, with an accurate evaluation for the whistle's feature characteristics. Furthermore, our analysis for real dolphin's whistles from 9 hours of data shows that, compared to the PamGuard benchmark software, ECV achieves high detection rate at a small cost of reduction in the false alarm rate.

Further work will improve the false alarm rate and the accuracy of the feature extraction by considering a non-casual spectral line analysis.

References

[1] Bossart, GD. "Marine mammals as sentinel species for oceans and human health". In: *Veterinary Pathology* 48.3 (2011), pp. 676–690.

[2] Wells, Randall S et al. "Bottlenose dolphins as marine ecosystem sentinels: developing a health monitoring system". In: *EcoHealth* 1.3 (2004), pp. 246–254.

[3] Oswald, Julie N, Barlow, Jay, and Norris, Thomas F. "Acoustic identification of nine delphinid species in the eastern tropical Pacific Ocean". In: *Marine mammal science* 19.1 (2003), pp. 20–037.

[4] Gillespie, Douglas et al. "PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localization of cetaceans." In: *The Journal of the Acoustical Society of America* 125 (May 2009), p. 2547.

[5] Oswald, Julie N. et al. "A tool for real-time acoustic species identification of delphinid whistles". In: *The Journal of the Acoustical Society of America* 122.1 (2007), pp. 587–595.

[6] Erbs, Florence, Elwen, Simon H, and Gridley, Tess. "Automatic classification of whistles from coastal dolphins of the southern African subregion". In: *The Journal of the Acoustical Society of America* 141.4 (2017), pp. 2489–2500.

[7] Esfahanian, Mahdi, Zhuang, Hanqi, and Erdol, Nurgun. "On contour-based classification of dolphin whistles by type". In: *Applied Acoustics* 76 (2014), pp. 274–279.

[8] Song, Hongjian et al. "An automatic identification algorithm of Yangtze finless porpoise". In: *Signal Processing, Communications and Computing (ICSPCC), 2015 IEEE International Conference on*. IEEE. 2015, pp. 1–5.

[9] Kohlsdorf, Daniel et al. "Probabilistic extraction and discovery of fundamental units in dolphin whistles". In: *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE. 2014, pp. 8242–8246.

[10] Rankin, Shannon et al. "Acoustic classification of dolphins in the California Current using whistles, echolocation clicks, and burst pulses". In: *Marine Mammal Science* 33.2 (2017), pp. 520–540.

[11] Janik, Vincent M. "Whistle matching in wild bottlenose dolphins (Tursiops truncatus)". In: *Science* 289.5483 (2000), pp. 1355–1357.

[12] Au, Whitlow WL. "History of dolphin biosonar research". In: *Acoust. Today* 11 (2015), pp. 10–17.

[13] Azevedo, Alexandre F et al. "Characteristics of whistles from resident bottlenose dolphins (Tursiops truncatus) in southern Brazil". In: *The Journal of the Acoustical Society of America* 121.5 (2007), pp. 2978–2983.

[14] Diamant, Roee. "Closed form analysis of the normalized matched filter with a test case for detection of underwater acoustic signals". In: *IEEE Access* 4 (2016), pp. 8225–8235.

[15] Forney, G David. "The viterbi algorithm". In: *Proceedings of the IEEE* 61.3 (1973), pp. 268–278.

[16] Piczak, Karol J. "Environmental sound classification with convolutional neural networks". In: *2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE. 2015, pp. 1–6.