Applied Acoustics 76 (2014) 274–279

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

On contour-based classification of dolphin whistles by type

Mahdi Esfahanian *, Hanqi Zhuang, Nurgun Erdol

Department of Electrical and Computer Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL 33431, USA

ARTICLE INFO

Article history:
Received 11 April 2013
Received in revised form 23 August 2013
Accepted 23 August 2013

Keywords:
Support vector machines
Fourier descriptor
Non-linear kernels
Pattern recognition
Dolphin whistles

ABSTRACT

Classification of cetacean vocalizations may help marine biologists study their behavioral context in different environments yet automatic classification of vocalizations for their information content has not been adequately addressed in the literature. Since classifier performance has a strong dependence on the extent to which features cluster, we, in this paper, explore the effect of two feature sets on two classifiers and assess their performance and computational complexity. We choose two feature sets that are exemplary of very different methods: The first set consists of Tempo-Frequency Parameters (TFPs) that are hand-picked to describe the spectral whistle contours. The second feature set embodies spectral information measured with the Fourier Descriptors (FD) commonly used in image processing for contour representation. The computed feature vectors are fed into the K-nearest neighbor (KNN) and Support Vector Machine (SVM) classification algorithms. The KNN in its basic form is a simple classifier that works well if feature clusters have clear margins and SVM uses a data dependent margin chosen for optimal performance. We argue that KNN serves to accentuate the effect of the feature sets and the SVM acts as the scientific process control. Experimental results show best results with the combination of the TFP feature extractor and the SVM classifier, suggesting a future research direction of developing non-linear kernels for SVM.

1. Introduction

The Southeast National Marine Renewable Energy Center within the College of Engineering and Computer Science at Florida Atlantic University is investigating the potential of harnessing ocean currents such as the Gulf Stream to generate base-load electricity, thereby making a unique contribution to a broadly diversified portfolio of renewable energy for the nation's future. Commercial turbines are being considered for installation in the Straits of Florida which is home to a large population of the Atlantic bottlenose dolphin (*Tursiops truncatus*), and several species of endangered whales are known to migrate through the region. Farms of underwater turbines with rotor diameters of 20 m may disrupt migration patterns and will likely act as fish aggregation devices, attracting predators such as the dolphins. The ability to classify vocalization patterns of these marine mammals for signs of distress and disruption, and distortion of their communication whistles will prove to be invaluable in predicting changes in their natural life cycle and health, and in developing turbine control measures to protect them.

Concern over anthropogenic activity in the oceans is not unique to the Florida coast. The near-shore location of the wave resource along the coast of Northern California and Oregon, for example, is known to be a migration route for the California gray whale (*Eschrichtius robustus*), and placement of extensive "farms" of wave harvesting equipment, including networks of cables, could present obstacles to the historical migration patterns of this protected species. Analysis of vocal patterns in the vicinity of planned installations could provide valuable input for behavioral studies that, in turn, could be used to mitigate possible disruptions to their behavior patterns.

Association of vocalizations with natural communication and behavioral messages has been the subject of research in bioacoustics for nearly three decades [1–3]. Recent studies matching behavioral activity to vocalizations have been reported in [4] by Panova et al. who investigated the relationship between different whistle types, pulsed tones, click series and noise vocalizations and six behavioral contexts by using the Kruskal–Wallis criterion and the Mann–Whitney U-test. Results of Fisher's exact test and the Pearson test to correlate dolphin whistles and 15 behavioral descriptions were reported by Ferrer-i-Cancho and McCowan [5]. Classification of dolphin whistle types is considered as a major step for biologists studying dolphin recognition and behavior [6]. Acoustic parameters of whistles were analyzed as potential

---

* Corresponding author. Tel.: +1 561 929 6394.

E-mail addresses: mesfahan@fau.edu (M. Esfahanian), zhuang@fau.edu (H. Zhuang), erdol@fau.edu (N. Erdol).

0003-682X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apacoust.2013.08.018

indicators of stress in bottlenose dolphins under various capture-release events and undisturbed conditions [7].

Dolphin whistles are narrow-band frequency-modulated sounds centered on a time-varying fundamental frequency commonly with several harmonic components [8]. They are believed to serve the purpose of conveying information and communication. Some species produce whistles with features that are sufficiently distinguishing to be categorized by function [9]. Dreher and Evans [10] argued that dolphins shared a function-specific repertoire of whistles, and each whistle was used in a particular behavioral context. The possibility of individual "dolphin identification" is also considered for some species based on the existence of features that transmit identity information that is independent of the whistle or the caller's location [11,12].

Time–frequency spectrograms are convenient media to explore and process dolphin vocalizations. They display long and short term data characteristics and capture their time-evolutionary spectral features. Treating them as images avails us of the use of archival image processing techniques to extract signature features that efficiently encode the distinguishing characteristics of the whistles. Feature vectors may be used as input to classifiers. This paper proposes using Fourier descriptors (FD) on the spectrogram as a non-parametric method of extracting contour features for automatic classification. Fourier descriptors are translation-, scale-, rotation-, and start-point invariant operators that encode shapes of two dimensional objects. We compare the performance of the FD feature vectors with the more direct but tedious method of Tempo-Frequency Parameters (TFPs) derived from traced contours. To test the efficacy of the FD features, we have chosen the K-Nearest Neighbor (KNN) and the Support Vector Machine (SVM) classifiers. We have chosen SVM because its class boundaries are data adaptive and are chosen to maximize the margin width between them. Increasing the margin width increases the measure of confidence in classifiers and reduces misclassification. KNN is a simple classifier that functions on majority voting. It is chosen as a "control" element: its output is used as a reference for judging the ability of a good set of features to cluster and the performance of SVM. To the authors' best knowledge, this is the first paper that reports the use of the FD as whistle contour feature vectors and the first paper to use the SVM to classify whistle types of dolphins and whales. That we have applied them to the classification of bottlenose dolphins' vocalizations by type does not preclude their application to classification by species or to other vocalizations displaying narrowband contours. We have tested combinations of the two feature extraction and two classification algorithms under consideration on the dolphin vocalization data and report the results and their analysis in this paper.

The steps involved in the automatic classification of dolphin whistles are typically detection, pre-processing, contour tracing, feature extraction and classification as described in Fig. 1. In the preprocessing stage we use two-dimensional band-pass filters to isolate the fundamental whistle and denoise the band-limited spectrogram. Under controlled environments of data acquisition and when real-time or even pseudo real-time processing are not the object, the easiest way to denoise is to extract the whistles by contour tracing in the spectrogram, and replace everything not on the contour with zeros. The isolation significantly improves the overall signal-to-noise ratio for detection purposes.

The remainder of the paper is organized as follows. Section 2 outlines methods used for pre-processing and tracing. The outputs are the fundamental components of dolphin whistles isolated from the remainder of the spectrograms after they have been time windowed and band-pass filtered. Section 3 describes algorithms used for extracting relevant features, and the methods used for classification of dolphin whistle types. Section 4 introduces procedures and equipment used for data collection, and presents experimental results along with discussions. The paper ends with concluding remarks.

2. Preliminaries: representation, denoising and extracting whistle contours

An effective method of displaying dolphin vocalizations is the time–frequency domain spectrogram. Spectrogram images of dolphin whistles show that each whistle is adequately modeled by a Fourier series with a time-varying fundamental frequency. Based on the data corpus used in the research, we have identified four types of whistle signatures in the time–frequency domain to comprise an exhaustive set to classify bottlenose dolphin whistles. A sample whistle from each class is shown in Fig. 1(clockwise from the top left): upswing, convex-up, convex-down and up-and-down [26]. We refer to them as Class 1 through Class 4, respectively. The frequency of the fundamental whistle varies in the range of 4–20 kHz over 250–500 ms with 3–5 harmonics (refer to Fig. 2) mimicking the frequency sweep at integer multiples of the fundamental. The harmonic relationship renders the higher harmonics redundant for characterization of the whistle so that we only need to work with the fundamental whistle. The proposed method is scalable to include more classes presented by a larger corpus.

Preprocessing steps depicted in Fig. 3 are used in order to extract contours of the fundamental whistle and include a cascade of high- and low-pass filters that limit the useful frequency range to 4–16 kHz. This is a generous bandwidth to contain all the fundamental whistles we have observed in our data set and to exclude low frequency environmental noise. Clicks and some parts of the second harmonic that remain in the pass-band are smoothed by a running-average filter. Reduced background noise and the absence of nearly all the non-fundamental harmonics are clearly visible in Fig. 4. The data are sampled at 80 kHz, organized into 1024-sample (~13 ms) frames, overlapping by 50%. Each frame is Hamming windowed, and the magnitude of its 4096-point Discrete Fourier Transform is organized in an array to be displayed as an image. It is assumed that the start and end points of dolphin whistles have been previously identified.

In most cases, fundamental frequency components have the highest intensity contours, as shown in Fig. 4; thus contour tracing can be obtained from its frequency location as a function of time. Spectral peak picking along each temporal frame is performed, yielding whistle contour as shown in dashed lines in Fig. 5. If the second harmonic competes, its frequency is measured and compared with the frequency of fundamental whistle. When the second harmonic is roughly an integer multiple of the fundamental frequency, it is considered as a peak lying on the second harmonic and consequently its corresponding fundamental frequency is interpolated using the neighborhood values on the fundamental whistle.

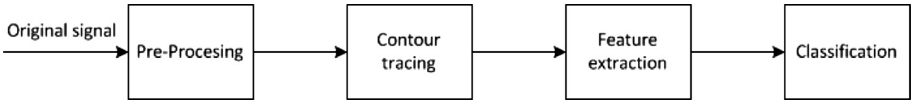

Fig. 1. Block diagram of the proposed approach to classify different types of dolphin whistles.