Applied Acoustics 119 (2017) 17–28

Contents lists available at ScienceDirect

Applied Acoustics

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

Marine mammal sound classification based on a parallel recognition model and octave analysis

Fernando Rubén González-Hernández, Luis Pastor Sánchez-Fernández *, Sergio Suárez-Guerra, Luis Alejandro Sánchez-Pérez

Centro de Investigación en Computación – Instituto Politécnico Nacional, Av. Juan de Dios Bátiz s/n, Nueva Industrial Vallejo, Gustavo A. Madero, México D.F. C.P. 07738, Mexico

ARTICLE INFO

Article history:
Received 6 October 2015
Received in revised form 15 November 2016
Accepted 29 November 2016

Keywords:
Underwater
Classification
Pattern recognition
Marine mammals
Acoustics

ABSTRACT

The ocean is full of a variety of sounds from natural, biological or anthropogenic sources. Listening to the animals sounds allows scientists to detect, identify, and locate different endangered species as well as listening to high intensity anthropogenic sources, which could harm the marine ecosystem. In this work, a new computational model for marine mammals classification is presented and validated with data from an online database. The feature extraction is performed using 1/6 octave analysis and the classification is carried out based on an independent ensemble methodology, where the outputs of four parallel feed forward neural networks are combined to classify eleven possible classes (seven marine mammals plus four additional classes). Unlike similar works, this paper considers multiple sounds emitted by each species such as whistles, calls and squeaks. The model demonstrated favorable performance reaching a classification rate of 90% at a low computational cost.

1. Introduction

In nature many animals use sound communication to exchange information. For instance, in the aquatic environment, marine mammals, including whales, depend on sound for both social interactions and to locate prey. For example, the use of passive acoustics to detect and classify species in-situ provides a means of identifying a species in their habitat, reveals their behavior as well as the population density. Automatic classification of marine mammal sounds is perhaps the most challenging task in the field of animal bioacoustics due to the unknown statistical signal properties, as well as the use of different recording systems and low signal to noise ratio (SNR) conditions, among others. Such discrepancies often lead to sub-optimal system performance.

This work evaluates different architectures for automatic classification of eleven marine mammal species found in the Gulf of Mexico, which is home to a high diversity of organisms. The model proposed herein could be useful to monitor, reduce, and avoid some human activities which occur in areas inhabited by protected species.

* Corresponding author.
E-mail addresses: f.er.moshi@hotmail.com (F.R. González-Hernández), lsanchez@cic.ipn.mx (L.P. Sánchez-Fernández), ssuarez@cic.ipn.mx (S. Suárez-Guerra), l.alejandro.2011@gmail.com (L.A. Sánchez-Pérez).

http://dx.doi.org/10.1016/j.apacoust.2016.11.016
0003-682X/

In this paper the sounds belonging to the following species are classified as:

1. Two mysticete cetaceans: *Minke Whale (Balaenoptera acutostrata)* and *Humpback Whale (Megaptera novaeangliae)*.
2. Five odontocete cetaceans: *Killer Whale (Orcinus orca)*, *False Killer Whale (Pseudorca crassidens)*, *Atlantic Spotted Dolphin (Stenella frontalis)*, *Common Bottlenose Dolphin (Tursiops truncatus)* and *Sperm Whale (Physeter macrocephalus)*.
3. One sirenia: *West Indian Manatee (Trichechus manatus)*.

This work considers multiple types of sounds emitted by each species including whistles, calls, squeaks, thumps, moans, and others. Therefore the identification of a given class is determined by the features extracted from any of these types of sounds, potentially reducing the time required to detect and classify the species. However, given the external conditions that the recording systems are exposed to, three more classes are included: natural (rain, bubbles, etc.), anthropogenic (vessel's engines) and unknown (Fig. 1).

Currently there are multiple observatories, such as the international program "Listen to the Deep Ocean Environment" led by the Laboratory of Applied Bioacoustics of the Technical University of Catalonia [28] or the ALOHA observatory, operated by the University of Hawaii [2]. Thus, the amount of recordings has grown exponentially, which demonstrates the necessity for applying

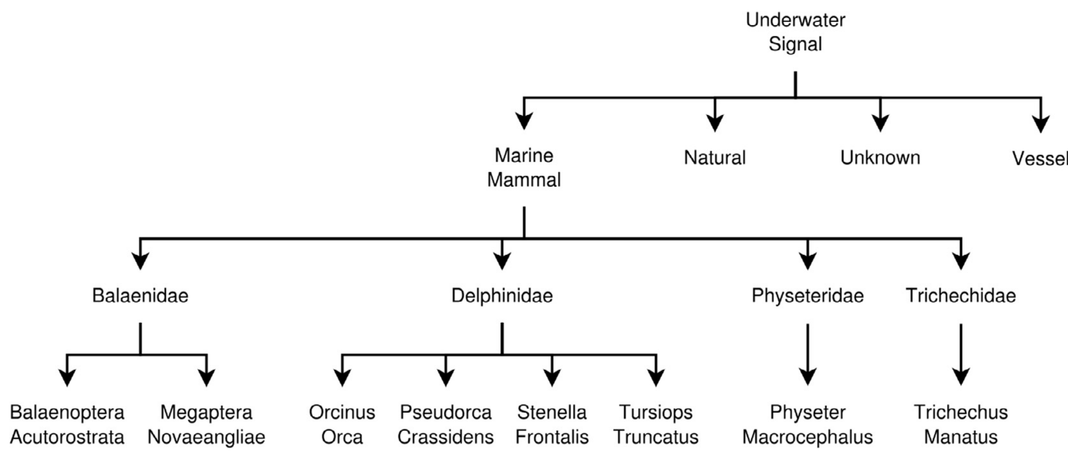

Fig. 1. Sources classification.

automatic methods to assess both on-site and off-site systems, in order to minimize manual interaction or supervision.

Underwater sounds are produced by a variety of natural sources, such as breaking waves, rain and marine life. It is also produced by a variety of man-made sources [23,4,18] such as ships and military sonars. Most sounds are relatively present everywhere in the ocean at all times. The background sound in the ocean is called ambient noise and is always present, whereby other sounds are only present during specific periods of time or places in the ocean.

Marine mammals [35] such as whales and dolphins produce sounds over a much wider frequency range, often outside the human hearing range. On the one hand, some large *Baleen whales* (mysticetes) produce sounds having frequencies lower than 10 Hz (below the human hearing range). On the other hand, dolphin echolocation clicks usually contain frequencies greater than 100 kHz (over the human hearing range). Other species also produce sounds such as the toadfish and drums, as well as some marine invertebrates like the snapping shrimp.

Passive acoustic classification is generally performed by the sonar operator. Presently, the final classification of a given sound is the responsibility of the operator [18,41]. However, by having these automatic classification systems the operator may make decisions with more confidence and being able using his skills mainly to analyze the most important or complex sounds.

Most of the works related to automatic classification of the noise produced by ships have dealt with features extracted in the frequency domain using Fast Fourier Transform (FFT) power spectrum [50,24,27], auto-regressive modeling [16,25] and wavelet transforms [16,10]. Regarding species detection, early automatic techniques make use of matched filters, Hidden Markov Models, and spectrogram cross-correlation [12]. These methods are later improved by using machine learning approaches such as feedforward neural network classifiers [34,39,13,33,32,40]. Other machine learning algorithms, such as classification and regression tree classifiers (CART), have also been implemented in recognizing contact calls made from the *North Atlantic Right Whale* [14,15]. Improvements over single recognition methods have been obtained by using an advanced technique that combines several recognition methods running in parallel [14,15,40].

Whales are widely studied mostly due to their unique communication capabilities. Abousleiman et al. [1] have developed an algorithm to pre-process the sound before applying a tree based hierarchical classifier. The main goal is to determine whether a *North Atlantic Right Whale* is present or not. They perform this binary classification by identifying a unique call made by the whale known as "*contact call*" or "*up-call*", achieving a success rate close to 85%. André et al. [3] and Zaugg et al. [53] detect cetacean emissions considering specific frequency bands, reaching a classification accuracy above 90%.

Existing schemes rely on the use of cepstral coefficients [9,42,37] as the input feature space used for capturing mostly pitch information on different vocalizations. Other approaches, such as auditory perception features, spectrograms, and frequency contours have been used as well [8,19,52,51,29].

PAMGuard is an open-source freely available suite of passive-acoustic monitoring software for marine mammals [20]. Oswald et al. [38] developed in ROCCA (currently incorporated in PAMGuard), which is an open source software package that measures 54 whistle contour features and is able to classify whistles of seven species and one genus: *Globicephala macrorhynchus*, *Pseudorca crassidens*, *Steno bredanensis*, *Stenella attenuata*, *Stenella coeruleoalba*, *Stenella longirostris*, *Tursiops truncatus*, and *Delphinus* species. The classifier deployed is based on Random Forest Analysis trained on 54 whistle contour features, yielding an overall successful classification score of 62%.

The proposed method is mainly built in two stages. First, an octave analysis is performed, which is widely used in acoustical analysis and audio signal processing. Although the signals to be classified are transient signals, these are still considered for feature extraction due to the frequency behavior. Second, a neural network model is used for identification of the eleven classes.

This paper is organized as follows. Section 2 gives a description of typical marine mammal signals, including a detailed explanation of the pre-processing, processing, and feature extraction process. Section 3 describes the neural model while experimental results are presented in Section 4. Finally, the conclusions are drawn in Section 5.

2. Material and methods

2.1. Spectral and temporal properties of marine mammal sounds

Social sounds of marine mammals are usually studied with a spectrographic analyzer, which determines the "instantaneous" frequency and relative amplitude of a signal as a time function, with the information usually plotted as a spectrogram. Many of the sounds emitted by marine mammals will have a pulse-like or burst-like property.

Sound emissions by *odontocetes* (toothed whales and dolphins) can be classified into two broad categories, frequency-varying continuous tonal sounds, referred to as whistles (see Fig. 2a) and broadband clicks [17], including burst pulse sounds (see Fig. 2b).