Detection and Tracking of Dolphin Vocalizations

Xanadu Halkias

Abstract

The field of audio engineering has predominantly been consumed by human centered applications. Many scientists have developed numerous successful algorithms that deal with important issues in speech and music processing. Automatic speech recognition systems, speech to text products or even music recommendation systems and simple transcription are just a few of the numerous research projects that have escaped research labs worldwide and have entered our everyday lives.

However, audio processing has also affected the field of animal bioacoustics. Many audio engineers are now using their knowledge to advance our understanding of the world that surrounds us and especially that of animals. This work falls into that category, where the principles of signal processing, communication theory and machine learning are used to analyze the clandestine world of marine mammals and specifically dolphins. Although substantial efforts have been made in the scientific community, there is still a need for the creation of an automatic and robust system that will allow for the detection and tracking of dolphin vocalizations.

In this work, Chapter 1 provides a short description on the physiology of dolphins focusing on how they produce and use sound for their survival. Chapter 2 is an overview of the most popular systems in the field for the analysis of marine mammal vocalizations and provides a framework for the task at hand.

Continuing, Chapter 3 describes the unique aspects of the data as well as the different features that are used in order to provide the proposed detection and extraction schemes.

In Chapter 4 several algorithms are proposed for the successful detection of dolphin calls in long recordings. Starting with the simple and widely used thresholding detectors, several advancements are proposed, based on the data, leading to more intricate classifiers like Support Vector Machines (SVM) that are known for their robustness.

Chapter 5 provides two systems for pitch extraction. The first system is based on a probabilistic framework and deals with the extraction of dolphin whistle calls while providing a first attempt on resolving simple overlaps. The second system assumes that the desired calls have already been detected and proceeds to identify the pitch for both whistle and burst calls using hierarchically driven Hidden Markov Models (HMM).

Finally, Chapter 6 presents the conclusions from the proposed methodologies and highlights both their advantages and disadvantages. Dolphins are intensely vocal mammals and although we are still unaware of the purpose or context of these vocalizations, I hope that by the end of this work we will have a better idea on how to uncover and approach their sounds, one of the few doors that allow us to enter their world.