A Theoretical Look at the Application of Shape Analysis to Acoustics and Signals

Alex Townsend, Peter Beerli, Anke Meyer-Baese, Dennis E. Slice

Florida State University, Department of Scientific Computing

Abstract

Over the past several years, a small community of researchers has arisen to develop and apply statistical Shape Analysis and Morphotoprties to the study of Acoustics and other Time-Varying signals. Here, we discuss the theoretical underpinnings of Shape Analysis and Morphotoprties as it applies to Acoustics and Signals. Ultimately, we show in this work that direct theoretical parallels can be drawn from the concepts and techniques of Morphotpirties to generate meaningful and valuable results in the field of Bioacoustics. We discuss how this theoretical framework differs from previous work and provide new insights into Bioacoustics and acoustic ecology. Using this theoretical framework, we build a pipeline for analyzing bioacoustic data in a Morphotoprtic context, and we show that the pipeline can be highly scalable to large problems. We also discuss algorithmic and implementational aspects of this pipeline including those for the extraction, alignment and analysis of individual pulses in a time-varying signal with pulse-like properties.

Introduction and Foundations

• Signals consist of a stream of data for which each data point is often correlated to its neighbors. Usually, they vary in response to some external stimulus or variable, such as the passage of time (2). Sound is one such signal.
• We assume here that all signals under consideration are of finite length and may only be sampled to an arbitrary but not infinite precision.

Definition 0.1. A Call is any set of \( n \) ordered pairs \( ((k_1, t_1), \ldots, (k_n, t_n)) \in \mathbb{R} \), where \( k_i \) represents an amplitude of a pressure wave in a fluid medium sampled at a specific time, \( t_i \), which represents a time such that \( t_1 < t_2 < \ldots < t_n \).

Definition 0.2. A Pulse is any subset, \( P \), of length \( m \) of a call \( C \), such that within a given window of size \( \omega \), noise mean \( \mu_X \) and noise variance \( \sigma_X \), then \( P \) is \( \{C_i \in C, \ i = (1, \ldots, m) \mid |\mu_X(C_i) \ | > |\mu_X(C_i + \omega) \ | \ \} \) \( \{C_i \in C, \ i = (1, \ldots, m) \mid |\sigma_X(C_i) \ | > |\sigma_X(C_i + \omega) \ | \} \) \)

Definition 0.3. A Shape is any set of \( n \)-dimensional points \( X_j, \ldots, X_k \) ordered by an index \( i \) such that \( i \) is strictly increasing. It is the information left in these points when rotation, translation and scale are removed (5).

Theorem 0.4. In the case of a linear two-dimensional shape (where \( x_1 < x_2 < \ldots < x_n \)), the definitions of Pulse and Shape are logically equivalent.

Pulse Extraction

Figure 1: An example how the pipeline finds and extracts individual pulses (red boxes). The call from which these pulses were generated was simulated using the SoundGen Package in R.

Pulse Extraction Theory

• The definition of the pulse leads nicely to a method to extract pulses from a call. For each call in your set of calls, loop over every neighborhood of size \( \omega \) with or without overlap, if the mean and variance exceed the tolerance, then add that to the current pulse being built. Once the mean and variance have fallen below the tolerance for enough window steps, finalize the pulse. If there is more than one pulse in a call, there will be multiple pulses in the list of pulses for each call. This will be \( O(N^2) \) in the worst case.

Efficiency and Scalability

• Single-Pulse runtime is \( O(N\log N) \) where \( k \) depends on the clustering algorithm chosen.
• GPA and the Pulse Extraction algorithm are nearly embarrassingly parallel.
• Both are capable of multi-level parallelism as well (multicore and multi-node).
• Thus, several terms in the runtime analysis can be divided by \( (MP) \) where \( M \) is the number of nodes available and \( P \) the number of processors.

Pulse Alignment

• In order to accurately measure the difference between \( n \) shapes, the shapes must be aligned such that each point in shape \( i \) is closer to it’s corresponding, homologous point in shape \( j \) than to any other non-homologous point in shape \( j \).
• Since we know that shapes and pulses are just different words for the same thing, the Generalized Procrustes Analysis method may be applied here to optimally align \( n \) pulses to one another (5). GPA is \( O(N^2\log N) \) in the worst case (1).

Analysis

• Once the pulses are maximally aligned as shown in Figure 2, then we may compute the shape distance between all pairs of aligned pulses, as would be done in Morphotoprties, thus generating an \( \mathbb{N} \times \mathbb{N} \) distance matrix where \( N \) is the number of aligned Pulses to be analyzed (4).
• We may use standard techniques to analyze this distance matrix such as Neighbor-Joining Clustering, in order to generate a tree of shape distances with known branch lengths such that dialect groupings may be seen (3). Neighbor-Joining is \( O(N^2) \) in the worst case (3). Most of the commonly used clustering methods are also \( O(N^2\log N) \) or \( O(N^3) \) (4).