Philippine Science High School
Main Campus

Real-time Identification of Simple and Extended Musical Chords using Artificial Neural Networks

Joachim Alfonso A. Navarro
Lesli Natasha A. Coronel

May 2019

INTRODUCTION

Background of the Study

Musical chords are groups of individual notes often played together and are an integral part of harmony in many forms of music. They are arranged in such a way that they follow certain harmonic rules that are recognized by humans as a response in the brain that is triggered when they are violated (Leino, Brattico, Tervaniemi, & Vurst, 2007). Because of this, chords are fundamental to the harmonic integrity of any musical work.

A musical chord is commonly identified and named by its root note and type (hereinafter referred to as the “chord type”). The root note serves as a reference point for the other notes which are played in the chord. These other notes are determined by the chord type. For example, a C major chord has “C” as the root note and “major” as the chord type. A major chord type includes the 1st (root), major 3rd, and 5th degrees of the typical Western major scale. Thus, the notes of a C major chord are C (1st degree), E major (3rd degree), and G (5th).

While “the general music-learning public places a high demand on chord-based representations of popular music” (Humphrey, Bello, & Cho n.d.), identifying the parameters of a certain chord takes certain skills obtained naturally or through special training (Zatorre, Perry, Beckett, Westbury, & Evans 1998), limiting the number of people who can correctly name chords by hearing them. Complete and accurate determination of these chords by hearing requires the use of both absolute and relative pitch, because chords utilize both an absolute reference point (root note) and a relative configuration of harmonies (chord type).

Absolute pitch is expressed when one can identify a musical note by hearing it, while relative pitch is shown when one can recognize the distances between musical notes (Zatorre, Perry, Beckett, Westbury, & Evans, 1998). While “most trained musicians” (Zatorre et al., 1998) exhibit a mastery of the latter, few have absolute pitch.

Absolute pitch is expressed in a low percentage of the human population and acquired through either favorable genes, music training at a young age, or both (Baharloo, Service, Risch, Gitschier, & Freimer, 2000).

Complete chord identification is thus a rare skill found in those with mastery of both absolute and relative pitch, even though chords play an important role in any musical work. An algorithm that automatically identifies chords from individual notes in real time would be a first step towards addressing this problem in the context of music education.

Artificial Neural Networks (ANNs) are computational models that use a layered structure of computational units called neurons in order to learn a certain task. Neurons in one layer are connected to all neurons in the next layer; each link has a strength called a “weight”, and each neuron has a value called an “activation” and another value called a “bias”, which allows its activation to be manipulated. Using a calculus-based process called *stochastic gradient descent* (SGD), the neural network manipulates its weights and biases until a certain parameter is minimized or maximized. For example, a neural network can be configured to minimize the value of an error function, which indicates the inaccuracy of the network. (Colina, Perez, & Paraan, 2017; Daniel, 2013; Nielsen, 2015; Sanderson, 2017).

Objectives of the Study

This study aims to develop a neural network that quickly and correctly identifies one-root musical chords of simple and extended chord types formed by playing more than two notes on a MIDI input device. Specifically, the program must identify common & extended chords and respond within 40 milliseconds, a standard determined by Greeff (2016) for use in live performance. The program must be implemented in programming languages that have MIDI input-output libraries such as *pygame* for Python and and neural network libraries such as *Keras* and *TensorFlow* to facilitate ease of implementation.

Significance of the Study

Current neural network implementations of chord identification such as that Perera and Kodithuwakku (2005) are limited to simpler chord types, and exploration of real-time chord identification using neural networks is limited to a few selected chord types: Perera and Kodithuwakku (2005), Osmalskyj, Embrechts, Pierard, and Van Droogenbroeck (2012), and Zhou and Lerch (2015) used two simple chord types only. Including extended chord types such as 7^{th}s, 9^{th}s, 11^{th}s, and their permutations would allow the application of chord identification on more styles of music, such as jazz and blues. The utilization of neural networks to identify extended chords from MIDI signals in real-time is largely unexplored and would provide useful data for future research.

A possible application of real-time chord identification is a program that displays the chord name of a set of MIDI notes passed to it through a physical device such as a MIDI controller. Such a program would be useful in the field of music education, where a low proportion of music students have absolute pitch (Gregersen, Kowalsky, Kohn, & Marvin, 1999) despite their demand for chordal representations of music (Humphrey, Bello, & Cho, n.d., par. 1). These allow said students to learn to identify the chords they are playing more quickly and accurately and help them develop their senses of relative and absolute pitch. They are also used in situations when musicians need to verify the chords they are playing for correctness, which usually happens when they are learning or composing a musical piece.

Scope and Limitations

The artificial neural network was configured to identify both non-extended ("common") and extended chords and aimed to display a result within a time limit of 40ms (Greeff, 2016). The following chord types were used: simple triads (major, minor, augmented, diminished), dominant extensions (7, 9, 11), major and minor extensions (M7, m7, M9, m9, M11, m11), suspended triads (sus2, sus4), major extensions with suspensions (M7sus2, M7sus4, M9sus2, M9sus4, M11sus2), dominant extensions with suspensions (7sus2, 7sus4, 9sus2, 9sus4, 11sus2), augmented and diminished extensions (aug7, dim7, ø7, aug9, dim9, aug11), and other extensions (mM7, mM9, M6, m6, M6(9), m6(9)).

MIDI notes were used as inputs to the neural network; audio datasets were thus not included in this study.