BRNO UNIVERSITY OF TECHNOLOGY FACULTY OF INFORMATION TECHNOLOGY

Excel@FIT 2023
http://excel.fit.vutbr.cz

# Web application for learning how to play the guitar

Michal Mikota*

**Abstract**
The tools for teaching music instruments have been with us for hundreds of years and, like everything else, have gradually adapted to the current times. However, a large number of these applications are too complex, confusing, and generally unfriendly to users. At the same time, practically no similar application has the feature of detecting the played tone by the player in real-time, which is very desirable and necessary for beginners and slightly advanced guitar players. This bachelor's thesis shows and explains the implementation of individual tools required for teaching using sound detection techniques and filtering out unwanted sound artifacts using the CREPE algorithm. This algorithm can detect and determine the played tone on the guitar even in noisy environments. The individual tools created within this application use this algorithm and, in combination with a simple user interface, provide the possibility of solving similar web applications.

*xmikot01@stud.fit.vutbr.cz, *Faculty of Information Technology, Brno University of Technology*

## 1. Introduction

Most guitar teaching applications do not adequately interact with the played note on the musical instrument, resulting in incomplete and non-comprehensive teaching. Applications like Yousician or Guitar Tricks are often too complicated for most players, with unfriendly user interfaces and no direct interaction with the player's musical instrument. However, their biggest advantage is the large amount of available materials for playing often created by large communities of guitar enthusiasts. Especially beginners may not realize when their instrument is not tuned correctly or when they play a wrong note while playing their favorite song.

Usage of FFT would be too complicated and not much worth for this kind of action. The main question was to find better way to detect dominante pitch from environment and use it properly in our future web application.

## 2. CREPE algorithm

This application uses the CREPE algorithm, which is based on a deep convolutional neural network operating directly on the input of the time domain. It is one of the most advanced and accurate algorithms for pitch detection, filtering out unwanted artifacts with an accuracy of between 95-98%. The algorithm takes 1024 samples from the time domain audio signal with a sampling frequency of 16 kHz as input. There are six convolutional layers whose result is a 2048-dimensional latent representation, which is connected to the output layer with a sigmoid activation corresponding to a 360-dimensional output vector. The resulting pitch estimate is calculated deterministically based on this output vector.

The ability to function well in noisy environments is crucial for many applications, such as analyzing speech on mobile devices or smart speakers, or performing live music. Among different algorithms, CREPE has shown to maintain the highest accuracy across all signal-to-noise ratio (SNR) levels for public noise and white noise. However, for pink noise at the highest SNR level, its performance is inferior to other algorithms. An exception is brown noise, where pYIN performs almost unaffected by noise. This is likely because brown noise has most of its energy at low frequencies, which the YIN algorithm, upon which pYIN is based, is particularly adept at handling. [**?**]

## 3. Application tools

This application included two main tools:

- Real-time tablature
- Guitar tuner with options for different tuning

### 3.1 Real-time tablature

Aligning musical input with the corresponding score or tracking the score in real-time poses a challenging problem for multimodal machine learning. Training a system that can solve this problem using sound and actual music notation (i.e. scanned or image-based sheet music) requires proper time synchronization between the sound and the coordinates of the notes on the sheet music images. In this work, this is one of the most problematic parts that requires suitable compromises between the number of input data on a given area and the least computationally demanding complexity during the detection of individual notes. In this case, the CREPE[1] algorithm is used in the implementation of the tuning tool.

For each song is necessary to prepare music notation or tablature. It's necessary to write it into an array that will contain objects with different types of notes[2]. An overview of the parameters needed for implementation includes:

- Pitch frequency
- Note type
- Duration of the sound
- Time point according to the overall duration of the composition

All of these parameters are then recalculated according to the playback speed setting (BPM settings) of the composition.

### 3.2 Guitar tuner

One of the most complex tools for implementation in terms of filtering and pitch detection as well as setting a user-friendly interface is definitely the built-in tuner.

In music or acoustics, a pitch is understood as a sound that has a periodic or at least approximately periodic waveform. It can be generated in various ways, but it always involves a certain kind of vibration (such as strings or vocal cords).

The main attributes of pitches in music are pitch, volume, duration, and timbre. The opposite of a pitch can be understood as noise (sound artifacts), which do not have a periodic waveform.

Pitch is determined by the frequency of vibrations. The higher the frequency, the higher the number of vibrations, and thus higher pitches. Pitch is closely related to frequency (although they are not identical terms). Frequency is an objective, scientific property of a pitch that can be measured. Pitch, on the other hand, is only subjectively perceived by the listener.

Each pitch is located at a certain frequency level. Using specific frequencies in music is unnecessarily complicated. The most prominent example is the so-called concert A, which has a frequency of 440 Hz.[3]

$$f_i = f_j \cdot 12\sqrt{2}$$

The sequence of pitches created by this formula gives us the chromatic scale. Instruments that are tuned in this way are most commonly used in Western music.

We denote individual pitch shifts with numbers, which we write in subscript. Lower pitches, which have a lower frequency, have an index of zero, and the highest pitches have an index of eight.

Two implementation options were considered for the decision:

- implementation of pitch detection using FFT (Fast Fourier Transform)
- implementation of pitch detection using neural networks

## 4. Conclusions

Application is mainly a good starting point for beginner guitar players. The advanced players will mainly benefit from the addition of chord playing options within the current tablature. Since chords are a more complex technical issue, I would like to implement this feature in future development stages of this application. Additionally, there is potential to further utilize the CREPE algorithm to develop other supporting tools such as pitch detection training tool.

## Acknowledgements

I would like to thank my supervisor Igor Szöke for his help with taking my own idea for bachelor's thesis. He has helped me with several problems such as signal processing or UX design for this web application. He also recommended me to be part of the this excel fit event.

## References

[1] Crepe. crepe. [online], 2018.

[2] Bc. Jakub Bajzík. RozpoznÁnÍ zvukovÝch udÁlostÍ pomocÍ hlubokÉho uČenÍ. [online], 2019.

[3] Lukáš Kotrbatý. Detekce dominantní frekvence. [online], 2019.