ISSN:0975-9646

Siddhant Gawsane et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6) , 2014, 6939-6941

Tonic Comprehension and Modeling of Analog Transmissions

An insight into the future of 'Smart Music'

Siddhant Gawsane[1], Koninika Patil[2]

[1] Department of Information Technology, SKNCOE
[2] Department of Computer Science, SKNCOE
Pune, India

***Abstract*—Tonic comprehension is the ability to interpret music as notation and therefore the ability to reproduce it on any given instrument of choice. However, this is in an art that takes long years of practice to develop, if at all; clearly it is the threshold that marks a novice musician from an expert. Smart Music is a mobile application that seeks to bridge this gap and thus ease the learning process for an amateur musician. It can allow anybody with even a little knowledge of music to play a song of their choice, on an instrument of their choice. The app correlates a musical note with the respective frequency and records all such frequencies that are being played at an instant. It takes as input the recording of the music piece and creates a log of frequencies played with respect to time. It then uses algorithms to prune out mismatching frequencies resulting from ambient noise, overtones or recording errors. These algorithms compare inputs with chord structures and popular musical overtures to determine music from noise. It finally re-creates the song in standard musical notation, or tablature, as the user chooses. Version 2 of the app will allow the user to pick an instrument that will be shown being played on the device's screen. Thus, a tuned instrument and a recording of the song are all that the user needs to play it. This paper aims to discuss various frequency analysis techniques and pattern matching algorithms used in the making of the Smart Music mobile app.**

***Keywords*—Music, Frequency analysis.**

I. Introduction

Plato once said, "I would teach children music, physics, and philosophy; but most importantly music, for the patterns in music and all the arts are the keys to learning". The app Smart Music eases and simplifies the process of learning music by bridging the most important hurdle, tonic comprehension. The primary objective of this app is to convert played music into written notation, thereby allowing the user to play it on an instrument. This involves an augmented analysis of the frequency distribution track.

The complete musical analysis involves a series of several steps. The first step is to identify all the instruments being played. This is done by the extraction and classification of features such as timbre, frequency and amplitude. Next, each singled out instrument track is thoroughly examined, and the notes being played are identified and converted into notations. Some of the methods already in place for this are Fourier analysis, spectrograms and the newest, scalograms, all of which have been discussed in more detail below. All these methods provide us with the time–frequency graphs which help to highlight the consistencies among frequency patterns of different instruments and also highlight the inconsistencies.

The possible usages of the frequency analysis techniques described herein are endless. The Smart Music app exploits one such possibility to teach the user the song that one has picked. It can be extended to analysis and correction of music, auto-tuning, differentiating patterns of music that are growing in popularity from the ones that are fading out, demonstrating the correlation between musical patterns and genres, documenting the evolution of popular music, and so forth. Needless to say, the techniques themselves must continue to evolve to overcome the limitations of dampening, background noise, overtones, etc. while the time–frequency analysis algorithms need to imbibe methods of self-learning to adapt to popular music trends.

II. Features Of Music

Before discussing the features, we need to understand their significance. All of the features of music need to be analyzed and extracted to complete the first stage of our project: Instrument Identification. Frequency is the lower level distinguishing feature which differentiates two notes based on their frequency. Timbre provides a deeper insight into the instrument being used and richness of sound. Each of these is discussed in avid detail below.

A. Frequency

The main feature of music is frequency, which decides which note is being played, i.e. the pitch. The piano key A4 has a frequency of 440Hz and B4 is at 466Hz. On the other hand, the frequency of A3 is 220Hz, thus demonstrating that frequency distinguishes between different notes as well as the octave of the note if the same note is being played. A note, however, always has a frequency, which is a multiple of its fundamental.

Here is a simple example which will immediately clarify the concepts of harmonics: when a string is plucked, if the entire vibrating string forms a single arc, it is the fundamental or first harmonic. Thus, the fundamental forms the lowest frequency which can occur on that length of string. If the string forms two separate vibrating arcs, it is the second harmonic and so on. Fourier series and harmonics are associated in the most very roots and we shall see how they are related further on.

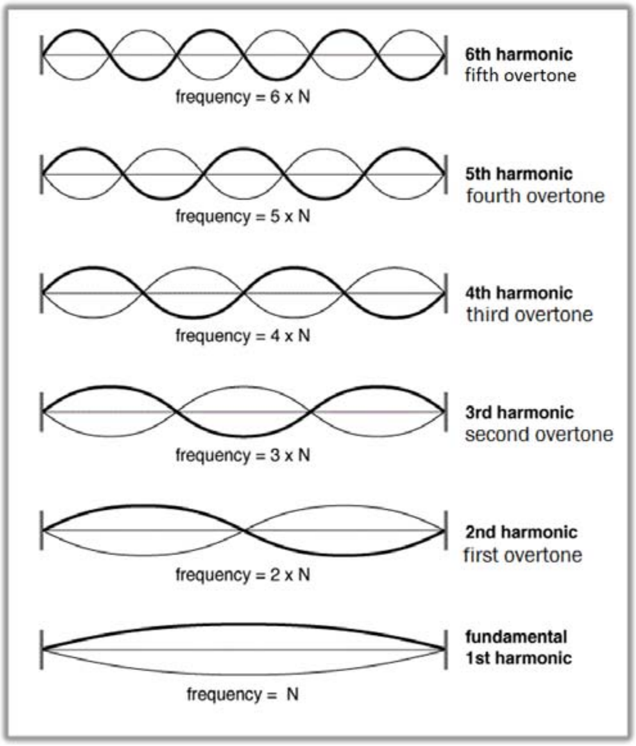

Fig. 1 Fundamentals, harmonics and overtones.

Music is made up of sounds, and sound is a wave. Therefore, it can be represented in the form of a sinusoidal equation. Consider $y(t) = A\sin 2\pi ft$, where f represents the frequency of sound and A is the amplitude.

B. Amplitude:

The amplitude pins a number on the strength or loudness of sound. It tells us the volume of sound, and it may vary over music and is also related to the mood of music. A silent piece is generally mapped to a sad song and loud music is attributed to cheerful music.

C. Timbre:

When a particular instrument plays a note, the music coming from it will be different from any other instrument. The music is not a pure tone but a sound emerging from the vibration or waves formed in different parts of the instrument. Some of the parts which vibrate to contribute to the quality of the sound are as follows: strings (sitar), membrane (table), hollow structure (drums), vibrating air column (flute), metallic body (Tasha). Also, special instruments such as violin have two parts striking against each other (bow on the strings), which produce a different quality of sound. Therefore, the guitar sounds different when played with a pick, nail or finger. Thus, there is a *source* which aggravates sound production and a *filter* which maintains the strength of the sound, both responsible for the sound quality. All these factors contribute to providing a unique sound to the instrument. The wave forms of different instruments are now understandably different and we can, on these grounds, differentiate between instruments.

However, another factor of the environment needs to be taken into consideration. All musical waves are produced in different environments, maybe in the solitude of one's house and a quiet melody on the piano, or a crowded mosh pit where the electric guitar is drowned out by the cheering of fans. So, we need to consider the noise factor which may create some interference on our only trump card: the wave.

Simple pattern recognition of wave forms can therefore be difficult. The solution is to pass the wave through a sound filter before analyzing it to get rid of the noise. A particular value is experimentally found which gives the least amount of loss of data and also gets rid of the noise. The wave is filtered based on its amplitude, which can be done since the volume of noise is less than that of the main track.

D. Duration:

As we discussed above, different filters or vibrating parts will prolong a sound for a different duration of time. A note played on the piano may be sustained for a longer time using the pedal. A lot depends on this duration with regard to analyzing the notes being played at any moment in time and shall be further discussed.

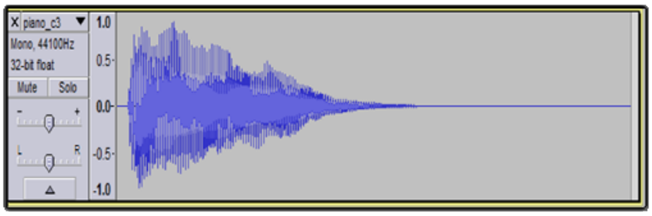

Fig. 1 Spectrogram of piano note C3

III. Tonic Comprehension

Now that we have identified the musical instrument, we separate the notes for this instrument and form a separate track consisting of only these notes. Once the individual instrument is segregated, we can carry out further analysis on this. There have been three ways devised for doing this.

The pattern of the notes and the frequencies are identified for finding out the notes being played. This frequency analysis is done in the following ways.

A. Fourier Series

A sound wave is best represented through the Fourier series. The Fourier series has the following equation:

$$f(t) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt) \qquad (1)$$

This equation can be expanded and re-written as follows:

$$f(t) = a_0 + (a_1 \cos t + b_1 \sin t) + (a_2 \cos t + b_2 \sin t) + \ldots \qquad (2)$$

The term $a_1\cos t + b_1\sin t$, which is the first non-constant term of the Fourier series, is the mathematical representation of the fundamental, $a_2\cos 2t + b_2\sin 2t$ is the first harmonic, $a_3\cos 3t + b_3\sin 3t$ the second and so on. This forms the harmonic series which has the mathematical form of the Fourier series. Harmonics are the integral multiples of the fundamental ($1f$, $2f$, $3f$…), whereas octaves only fall on the power of 2 ($2f$, $4f$, $8f$…). It also efficiently tells us the frequency and amplitude for each note being played by this representation. Thus, it also determines the tonal quality of the note.

Coming back to the musical notes, A1 (55Hz) is an octave over A0 (27.50Hz) and is hence a multiple of the fundamental frequency at A0. But the first harmonic for A0 is twice the frequency at A0, which here coincides with A0. The second harmonic however is not A2 (110Hz) but thrice 27.50Hz, which is E2 (82.5). Harmonics always sound good to the ear and; hence, many a time the first and fifth notes (5^{th} note being the 3^{rd} harmonic) always sound good when played together and form the integral part of a chord.

The Fourier series has a use other than identifying the overtone. It can be used to model sound, or *Produce* a wave having particular sound characteristics. Thus, Fourier series is very important for further advances in musical analysis. We shall further see how the spectrogram and scalogram pickup where Fourier series analysis left off.

B. Spectrogram

Spectrogram analysis is what we will mainly be using for identifying the frequency content or the actual notes being played. Spectrogram is especially required for polyphonic tonal comprehension as the Fourier series can only deal with so much as one note played per moment. To do this, we need to magnify every portion of the frequency pattern obtained by Fourier transform of the track. For this, we divide the frequency–time graph into windows at regular intervals and each window is analyzed in greater detail.

We need to build input and output pairs formed by the spectra of the sound produced by a source. Consider a given instant t_i. The input is $\{S(f, t_{i+j})\}$ for $j \in [-m, n]$ where f is the frequency at instant i, m and n are the number of windows considered before and after the central time, t_i. The output is a vector consisting of all the notes which are activated at instance t_i.

We need to consider the windows before an instant as the sound of a note played earlier carries over to our current instant time frame. Whether the note carrying forward is active or not is decided by a threshold amplitude value. All of the frequencies clearing this threshold are considered to be active and are added to the vector $v(t_i)$. So, for each instance t_i, there are k notes which are active if they have cleared the threshold amplitude and $v(t_i)$ will contain these for each instance. This vector is essentially in the binary form and 0 represents that the note is inactive and 1 that it is active. Once we have the activated notes vector, we can successfully find out the notation for the entire sound track.

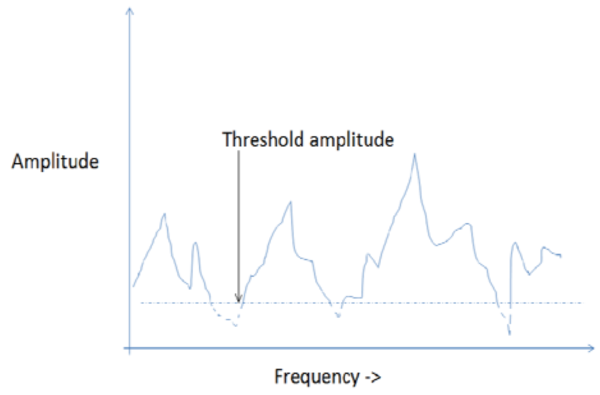

Fig. 3 A simple waveform is shown. Frequencies lying below the threshold amplitude are not considered as they could be noise resulting from inactive notes. Only the frequencies having amplitude greater than the threshold are considered active.

C. Scalogram

Scalograms and spectrograms are almost the same. Scalogram differs from a spectrogram in just the sense that it divides the frequency pattern into windows of unequal sizes, depending on the complexity or the greater number of colliding notes. Thus, a denser region in the frequency–time graph will have smaller windows, while a sparser region will have a larger size window as the notes will be easily distinguishable. These expansions and contractions will require a logarithmic scale for the frequency axis. Using the logarithmic scale, we can also deal with smaller values as the frequencies of the notes increases exponentially with each octave. Thus, scalograms deal with more delicacy, than what spectrograms were dealing with.

IV. UTILIZING THESE METHODS

Using the above-mentioned methods, we are able to separate the tracks of music and map the frequency of each note on a virtual music interface as given below:

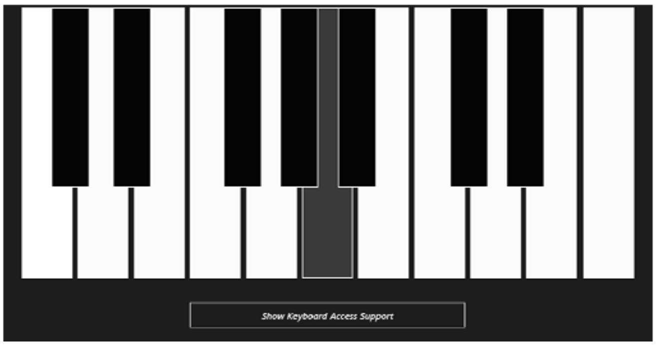

Fig. 4 A virtual piano interface

Such mapping is done for each and every musical instrument and simultaneously as the exact notes are found from the scalogram, even the textual notations can be charted out.

V. CONCLUSION

Programmatic tonic comprehension is all set to revolutionize the teaching and learning of music. Although a human teacher can never really be replaced, technology can surely reduce the teacher's burden and allow the automation of routine tasks. It can also facilitate music enthusiasts who prefer a non-classroom oriented approach to learning.

Of course this technology can also be used for a multitude of other applications. It can be used, for example, for different types of scientific analyses and rectifications of music. Auto tuning, though often criticized, is a direct application of this technology. Pitch adjustments of music in a karaoke-like scenario is another possibility. The list is endless.

REFERENCES

[1] J. Alm and J. S. Walker, Time-*frequency analysis of musical instruments*. SIAM Review, 44, 457–476, 2002
[2] J. S. Walker, *Music Analysis And Processing With Audacity And FAWAV,* Documentation for Audacity and FAWAV, 2002
[3] K. W. Berger, Some Factors in *the Recognition of Timbre,* J. Acoust. Soc. Am., pp. 396-406, 1964
[4] J. K. *Hammond, Mathematics of Music*, UW-L Journal of Undergraduate Research , XIV, 2011