Musical Timber Assessment Using MFCCs and Total Harmonic Distortion in Matlab

Hamed Sadeghi[1], MohammadMehdi Eskandari Jajarm[2]

Department of Electronic and Communication engineering, Tenaga National Universiti

Abstract: *Colour of music known as timber is a significant feature of sound which represents the character of it. If a musical instrument is perceived differently from other type while playing the same note, it is all because of their timber. The most important impact takes place when the music is recorded and the colour is changed. When someone listens to a piece of music he/she can identify whether the music is live or recorded. So if recording procedure and environment recording are set up in a way that they could produce the best timber, then it can be claimed that the timber of music is natural after recording. Timber extraction using Mel Frequency Cepstral Coefficients (MFCCs) method is done in this project to show that visual extraction can do nothing with differentiations between timbers and it is not possible to comment if the music is natural enough after recording or not by only looking at the timber. Then Total Harmonic Distortion calculation is employed in MFCCs to see how timber is distorted after recording, and then a group of normal and expert musicians are asked to comment about those pieces of music to make a comparison and reach an approach for musical timber assessment. Finally an estimation of timber checking is suggested which can be used by sound engineers to check if the recording environment and the musical recording equipment would result in natural timber after recording.*

Keyword: ditortion, enviroment, harmonics, musical instrument, Timber

1. Introduction

Digitalization of music and MIDI renderings of musical scores affect the quality of music in different ways and executing modern digital music synthesizers reproduce the pitches, note timings and even musical instrument timbres.The question is why they are often uninteresting and emotionless in comparison to the live performances of musical artists! Moreover, what are the alterations and shaping of musical pieces by the experienced performer that provide music's sensational influence [1].

1.1 Music Recording Environment

The main reason to do this project is to identify if the recording environment provides acceptable timber for the music. Actually there are different approaches for timber extraction; however in this project a method is used that results in values for each timber, so that comparison between timbers would be easier. Fixed audio recording system is employed and different recording environments have been considered as variants to check out how recording condition may change the timber of music. What is attempted in this project is to chose the most common recording environments in which most musicians prefer to record their music in and because of wide range of instruments around the world and diverse auditory sensation of music color, specific kinds of instruments called 'string instruments' have been analyzed.

Timber extraction does not solely assist in finding out whether the color of sound is natural or not, because the extracted timber is unique for that relevant sound. So timber values are analyzed over all conditions to see how timber of same music reacts in the different recording environments. And that timber values are acceptable or not, would be determined only if compared with musicians' comments about recorded samples. Therefore a range of acceptable value for timber can be reached to identify whether the environment provides the recorded music with natural color or not.

1.2 Timber Assesment

This project examines how natural the timber of recorded music is, so that the listener can perceive the music in a way that recorded one is very similar to the live one. Distinguishing timber is one of the most challenging issues in the music community. This is also an issue for those composers who employ some traditional/folk instruments in a different way by changing how they are played or applying physical modifications. In music perception research, Timber is that aspect of sound engineering that has not been understood precisely. Recently, pitch and rhythm have been investigated by many researches and they have had a great deal of recognition, but timber quality identification for string traditional instruments is an area that is ripe for examination.

Presenting a method of timber quality identification can help musicians to choose their recording environment regarding their timber quality considerations, so that the best tone quality would result. Timber quality identification can also be used in human speech modification in communication system to preserve and transmit human voice as natural as possible, however this project does not go through human voice identification and it is specifically done for musical proposes

2. Literature Survey

Based on previous sections describing timbers, listeners perceive different timbers for the sorts of musical instruments employed as more or less related. These relations in terms of spatial representations can be summed up. They are sometimes called timber spaces. Studies were done to distinguish if this kind of information can be used to suggest methods in which music sorting or music

organization is done by utilizing sequence of timbers. Some experimental studies [2] to test whether it is possible to get transpositions in timber space by creating analogies of the structure like ' A is to B as C is to D, where A, B, C, and D are different timbers. Listeners heard four possible solutions for each analogy problem and they were asked to comment about how good the analogies were for each four samples. According to predictions the last timber was supposed to be closest to the ideal point.

2.1 Musical Instrument Extraction through timber classification

Auditory system of human ear has the ability to differentiate between different of musical colors. Human ear can extract an specific instrument using the knowledge of tone quality or timber of the sound. Many audio features have been used lately in studies to attain automatic timber recondition system using different methods like machine learning algorithm. Recognizing audio features can be a challenging problem cause it is a continuous task, unlike discrete data. Due to reach precise predictions, recognition needs to be synchronized as in time frame. And in order to reach a suitable set up for the recognition, musical characteristics and timber are needed to be more explored.

Timber can be a series of individuals' comments about a sound that is unique in frequency or the amplitude. As explained before, timber is known as tone of sounds. There is no considerable attributes to timber, unlike frequency which is attributed to pitch or amplitude to loudness. This constraint makes the definition of timber to change and immanent.
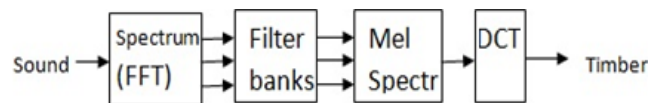
2.2 Timber extraction using MFCCs

Music industry landscape has changed completely by digital technology and internet. The large number of accessibility of music has let users to store and share thousands of files including different sorts of music with different sizes on their computer's hard disk, portable media player, mobile phone and other devices. There is a need for new applications for browsing, managing, discovering as well as generating playlists for users, by given the large music allocations available. The study of Music Information Retrieval (MIR) leads to address these challenges by using content-based techniques for executing tasks such as audio music similarity estimation and genre classification to distinguish between different sorts of music type [3].

3. Research Elaborations

This projects suggest timber extraction using MFCCs to comment about reording enviroment. In order to do timber extraction, sample of music has to be analysed in time and Mel domain which would be explained in following sections. Total Harmonic Distortion analysis would be explained seprately and then employed in extracted timber to conclude with a new method.

## 3.1 Overview of MFCCs extraction method in this project

Sound is expected to be inserted to the system as an input and there would be timber as output of the system. Fig 3.1 illustrates how timber extraction using MFCCs is done in this project.

Figure 3.1: Block diagram of timber extration using MFCCs

The process of timber extraction using MFCCs shown in fig 3.1 in this project is done by a toolbox in MATLAB called MIR Toolbox.

Timber extraction in this project is based on MFCCs. Fig 3.1 shows the diagram of operations. First, the audio sequence is loaded and converted into the spectral domain, using functions that MIRtoolbox has provided for this purpose. The spectra using filterbaks are converted from the frequency domain to the Mel-scale domain: the frequencies are rearranged into 40 frequency bands called Mel-band. The envelope of the Mel-scale spectrum is illustrated with the MFCCs, which are attained by applying the Discrete Cosine Transform to the Melscale spectrum. Because of repetition and non important regions in Melscale [4], usually only a restricted number of them (for instance in this project the 13 first ones) are selected. The computation would be concluded in a window through the audio signal, resulting in a series of MFCC vectors. Figures 3.2.a, 3.2.b and 3.2.c show an example of audio sample in three phases including audio as frequency spectrum, melscale spectrum and MFCCs respectively.
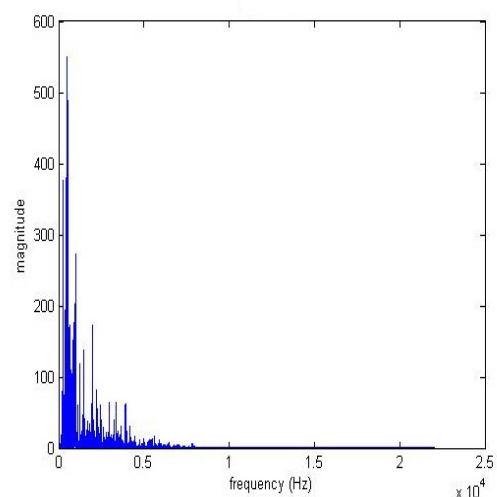
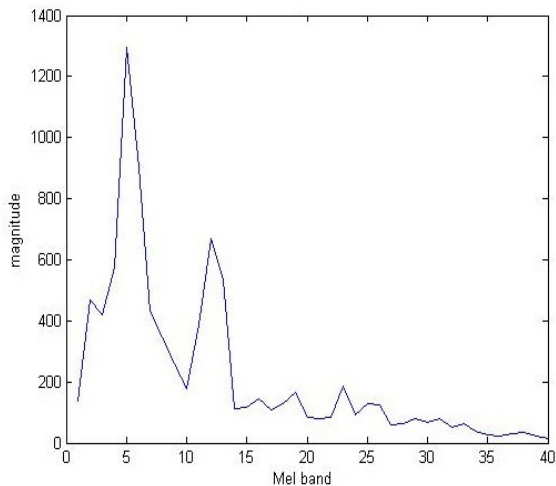**Figure 3.2 (a):** Mel spectrum of a sample

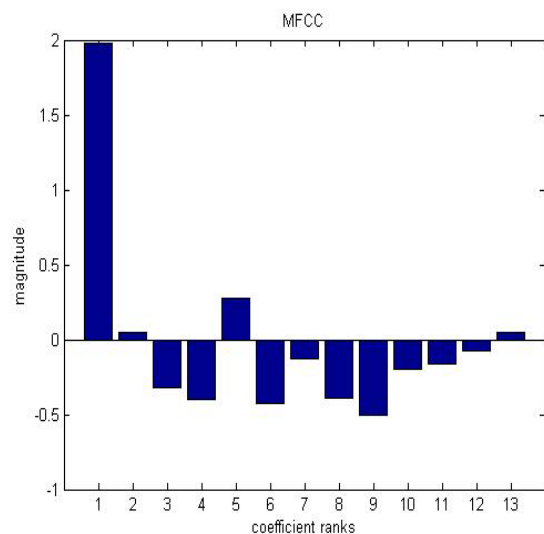**Figure 2.3 (b):** Audio spectrum of a sample

Figure 3.2 (c): MFCCs of a sample

3.2 Total Harmonic Distortion in MFCCs

The main concept of THD is used in this project after extracting MFCCs and it is called THDm. Since the first vector represents fundamental frequency and the rest 12 vectors are harmonics of the audio, then finding Total Harmonic distortion in MFCCs would make completely sense. This factor has never been used in previous researches and in this project THDm is applied to reach an absolute value for total harmonic distortion.

MFCCs represents 13 vectors in which the first one is fundamental frequency and the rest 12 ones are other harmonics which convey specific colour of music. As explained before timber or colour of music refers to harmonics of audio and the matter of differences between colours of two samples of music relates to their harmonics. However what this project seeks for, is to distinguish whether the environment can present a good timber after the music is recorded.

4. Recordings and Results

Here in this project a piece of recorded music is analyzed in 7 conditions virtually. The reason why simulation is done here is that an exact piece of music is considered to be play in different conditions. Therefore a musician was asked to play a piece of music and then imported in different virtual conditions. The recording sample is done in a well-prepared acoustic room providing all good recording conditions. And different virtual conditions are provided with Cubase LE software. These different conditions are considered in ways that are similar to recording environments that a musician may record his/her music in.

Seven musicans are asked to comment about each sample. And a comparion is made between musicians' comments and MFCCs values to reach an aproach to assess how natural the timber after recoding is.

4.1 What is Expected to Achieve From Results

Colour of music known as Timber is extracted from each sample of music (played by different instruments) in different environments and a value of THDm-measured mathematically by MATLAB-is dedicated to each condition. On the other hand all these samples are presented to 5 musicians with diverse proficiency in each instrument to comment how naturally each music is perceived after being recorded in all those environments. Each musician is chosen with different proficiency of skill in instrument in order to have different interests and even a specific musician is not familiar with a specific kind of instrument so that he/she can be a normal listener in that case. What expected here is that musicians' comments match the mathematically extracted THDm for the samples.

5. Conclusion

Different environment aspects influence the music being recorded in a different way. Terms like echoes and reverberation in recording auditory environments would have a number of dramatic effects, comprising increasing the realism of the display, distance perception improvement, presenting effective information about the room, and degrading directional accuracy. Relatively some issues remained unknown, something like, which aspects of reverberation are most vital for each of these perceptual results.

These results have a effective approaches for the design of effective, efficient acoustic room simulators, pointing to the need to take into account how different features like reverberation, echoes, mic distance or other features influence perception. More specifically, further studies are required to examine how accurately room reflection patterns should be simulated in a virtual environment to reach accurate distance perception as well as realism, or how other recording features must be simulated in virtual environment to best simulate the real recording, while some work addressed these issues.

Most sound engineers overlooked the acoustics as a key to capturing great recordings. If an instrument is placed anywhere in the room and put a microphone in front of it, that is basically like rolling the dice and hoping a good sound comes up. If there is a little more concern about how an instrument is placed in a recording space, then

Paper ID: OCT141320

2626

significantly better results will be achieved, with much less effort, even if inexpensive recording equipment is being used.

The way the right acoustic environment is chosen, and the way that immediate space around the instrument is treated, can be unique to each instrument and the sound which is trying to get achieved. By timber detection and calculating THDm, naturalness of the timber in acoustic area can be stated.

References

[1] N.H. Fletcher, *The Physics of Musical Instruments*, (New York, Springer, 1998)
[2] D, E., *Perception of timberal analogies,* Paris, IRCAM, 1978
[3] Leon, F.D.a.K.M., *Enhancing timber model using MFCC and its time derivatives for music similarity.* Southampton, Europian Signal Processing Conference, 2012
[4] Cook, G.T.a.P., *Multifeature audio segmentation*, IEEE Workshop on Applications of Signal Processing, 1999

Paper ID: OCT141320   2627