NOVEL MP3: THE MUSIC WITH SENSES

Kamna Singh
Department of Computer Science and Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, India

***Abstract*—The field of computer science known as artificial intelligence (AI) is dedicated to building machines that exhibit behaviours deemed intelligent by humans. Since the beginning of time, people have been fascinated by the idea of creating intelligent robots. With the development of AI programming techniques, this dream is now starting to come true. Researchers are developing ways to imbue computers with human-like intelligence. Systems that can imitate human thought processes, comprehend speech, gestures, fingerprints, and even facial features are being developed. These systems can also outperform the world's top chess player, among many other previously unachievable tasks. Humans are inherently musical, which is one of the main characteristics that set us apart from other species. Although most people equate music with emotional expression, it also makes sense that the intellect is a major component of musical endeavours. The interaction between these two components is a topic of study in many scientific domains, including neuroscience.**

***Keywords*—MP3 track, equalizer, gain level**

I. INTRODUCTION

First of all, to name just a few Machines, cognitive sciences and artificial intelligence (AI) are fields in which music is unquestionably one of the most fascinating applications of human intelligence. Through examining simulations of this task, scholars endeavour to unravel the enigmas inherent in both music and intellect. However, from a practical standpoint, the ultimate objective of research on AI and music is to teach computers to perform like expert musicians. In addition to highly specialist skills like composition, analysis, improvisation, and instrument playing, skilled musicians should be able to execute less specialized tasks like reading concert reviews in the newspaper and interacting with other musicians. In this scenario, the music machine would require a rudimentary comprehension of human social issues, including grief and happiness. The ability to perform a variety of such tasks is currently being searched for ways to integrate them into research projects. For example, mechanisms from music analysis systems are combined with systems for composition to enable computers to compose music autonomously in the style of pieces that have been analysed, filtering different aspects of the music and echo effects, adjusting speed and effects, etc. The processing of an auditory signal, or sound, is known as digital audio processing. It describes the process of sending and receiving sound that has been digitally preserved. A digital audio signal that has been informationally encoded is called digital audio. At a specific sampling rate and bit resolution, an analog signal is transformed into a digital signal. When it comes to digital audio processing, the main focus is usually on analysing the signal's audibility mathematically. An audio stream, for instance, can be altered for many reasons and regulated inside the auditory domain. Psychological characteristics have a major role in determining which portions of the signal are perceived as being heard and which are not, in addition to the physiology of the human hearing system. Psychoacoustics is the field that analyses these features. Storage, level compression, data compression, transmission, and enhancement (such as equalization, filtering, noise elimination, echo or reverb addition, etc.) are among the processing techniques and application areas.

Because the MP3 format allows sound files to be compressed while maintaining a high degree of quality and making them much easier to keep or send online, it has become the most widely used audio format nowadays. The majority of MP3 players are equipped with an inbuilt graphic equalizer, a high-fidelity audio control that lets the user separately adjust and view several different frequency bands. Equalization of audio is particularly beneficial in improving sound quality and limiting the physical extremes that might otherwise result from recording analog records. Higher-end and mid-range stereophonic sound systems for home usage frequently have graphic equalizers. Every time a user listens to a sound track, these equalizer settings need to be adjusted. This can cause a number of issues, including the deterioration of high-quality speakers due to abrupt changes in the track without corresponding equalizer settings, car accidents caused by the driver's distraction from manually adjusting these settings frequently, listener annoyance from constantly adjusting the settings, and more [1].

II. REVIEW OF THE LITERATURE

There is a plethora of material available on the present MP3 format and its decoding, despite the fact that there is very little study directly applicable to the desired music format or its decoding. Articles about the various equalization techniques that can be used on an audio stream are also available.

MPEG-1 Audio Layer 3, or MP3, is a well-liked lossy compression and digital audio encoding format. Its goal is to significantly reduce the amount of data needed to represent audio while maintaining what most listeners consider to be an accurate reproduction of the original, uncompressed audio. The

audio-specific compression format is called MP3. Compared to simple techniques, it offers a representation of pulse-code modulation-encoded audio in a substantially smaller amount of space by employing psychoacoustic models to filter out sounds that are less detectable to human ears and efficiently store the remaining information. A variety of bit rates can be used to compress MP3 audio, offering a range of trade-offs between sound quality and data size. As a general rule, the higher the bit rate employed, the higher the quality during playback since more information from the original sound file is contained. The possible sampling frequencies in MPEG-1 Layer 3 are 32, 44.1, and 48 kHz, and the available bit rates are 32, 40, 48, 56, 64, 80, 96, 112, 128, 160, 192, 224, 256, and 320 kbit/s [5].

Since MP3 files contain audio separated into frames, each of which has a unique bitrate, the bit rate can be dynamically altered while the file is being encoded. This method further improves quality and reduces storage space by using more bits for parts of the sound with higher dynamics (more movement in the sound) and fewer bits for areas with lower dynamics.

2.1. Graphic Equalizer: A graphic equalizer is a set of filters, each with a fixed center frequency that cannot be changed. The only control you have is the amount of boost or cut in each frequency band. This boost or cut is most often controlled with sliders. This interface is pretty intuitive because the frequency response of the equalizer resembles the positions of the sliders themselves. The sliders are a graphic representation of the frequency response. A graphic equalizer uses a set of bandpass filters that are designed to completely isolate certain frequency bands. Each filter in the graphic equalizer has the same input. Their job is to only allow a small band of frequencies through. The filters are arranged in parallel. For each filter that is added in series, its phase response is added to the phase response of the other filters. The phase response also reveals how the filter actually delays the signal.

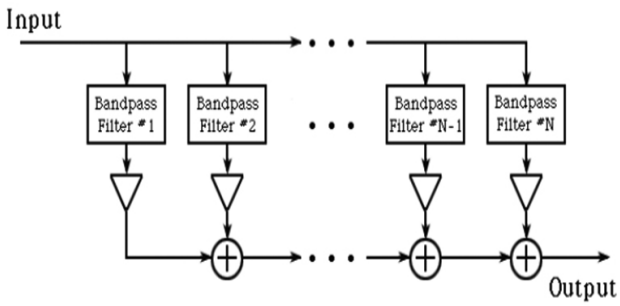

Figure 1. Graphic Equalizer

The MP3 header and MP3 data are contained in each of the several MP3 frames that make up an MP3 file. The real audio payload consists of MP3 data. The figure demonstrates that a sync word, which is used to determine the start of a legitimate frame, is a part of the MP3 header. MPEG-1 Audio Layer 3, or MP3, is the result of the following: two bits that indicate that layer 3 is being utilized, and a bit that indicates this is the MPEG standard. Following this, the values will vary based on the MP3 file. The ID3 metadata found in the majority of MP3 files today either comes before or after the MP3 frames; the figure illustrates this as well.

2.1.1 The equalization of audio signals is explained in this article. Tone controls, parametric equalizations, and graphic equalizations are covered. The practice of increasing or decreasing specific frequency components in a signal is known as equalization, or EQ. When recording, equalization is a crucial tool for enhancing the sound of an instrument. Beautiful music is authentically replicated for listeners in their rooms, giving them the impression that they are at a live performance. Tone controls: Probably the most widely used equalization mechanism, tone controls are present on the majority of stereo systems. They offer a quick and simple technique to partially offset the room's acoustics and customize the sound to your preferences. The lowpass shelving filter is controlled by the "bass," and the high pass shelving filter by the "treble." A signal is amplified by a gain larger than one, and reduced by a gain smaller than one. Shelving filters: only adjust or cut a piece, leaving the remainder untouched. 'In-between' frequencies are affected by'mid' controls, like the 3-band equalizers frequently seen on mixers. Typically, it doesn't try to isolate certain frequencies, but instead selectively increase or decrease a specific range of frequencies without affecting other frequency bands. Graphic Equalizers: A graphic equalizer consists of a collection of filters, each having an unchangeable fixed center frequency. You are only able to adjust how much boost or cut is applied to each frequency band. Most typically, sliders are used to manage this cut or boost. Because the equalizer's frequency response mimics the slider placements, this interface is rather intuitive [2]. The frequency response is shown graphically by the sliders. A series of bandpass filters intended to fully isolate specific frequency bands are used in graphic equalizers. The input of every filter in the graphic equalizer is the same. Their function is to restrict the range of frequencies that can pass through. In parallel arrangement are the filters. The phase response of each filter applied in series is added to the phase response of the preceding filters. The actual signal delay caused by the filter is clearly visible in the phase response [3].

2.2 Parametric Equalizers: A single parametric equalizer lets you adjust the bandwidth and center frequency in addition to the boost or cut. A parametric EQ with a lot of cut, often known as a notch filter, can be placed precisely at the frequency where the feedback is occurring in order to negate it. Theoretically speaking, shelving and peaking filters are just particular kinds of parametric equalizers; nonetheless, the majority of commercial devices do have parameter limitations.

2.3 A Whole MP3 Decoder on a Chip Hugo Hedberg, Thomas Lenart, and Henrik Svensson - CCCD, Department of Electroscience, University, 118, SE-221 00 Lund, Sweden This paper presents two distinct hardware decoder implementations, one for FPGA and the other for ASIC. Along with implementation details for the two decoders, it also offers basic information regarding the MP3 format. The creation was completed in a project-based learning environment [4].

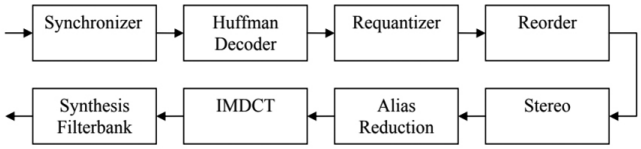

Figure 2: The MP3 Decoding Process

The suggested method calls for the equalizer settings to be adjusted for each sound track each time it is listened to. This causes a number of issues, including the deterioration of high-end speakers when tracks abruptly change without matching equalizer settings, car accidents caused by drivers getting distracted when they manually adjust these settings frequently, and listener annoyance from tuning the settings repeatedly[5]. If the equalization settings were saved in the audio format itself along with each sound file, these issues may be resolved. The music players could then decode the settings and set them automatically before playing each track. MP3 players can be configured to decode embedded equalization settings and adjust them properly before playing any audio. These settings can be found in the tags of the MP3 format. To allow for the tuning of different frequency vocals with varied equalizations, it is suggested that the equalizer settings for the recording's individual components be incorporated in the new format rather than the equalizer settings for the entire file. This would be helpful for songs sung by performers whose voice frequencies differ significantly as well as for songs including both male and female vocalists. Because the frequency of a woman's voice is higher than that of a man, different musical instrument settings are required to improve the audio response. One might therefore have more control over the sound track by using the time-varying equalizations.

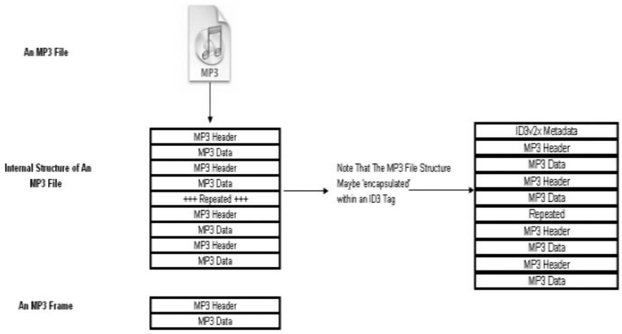

Figure 3: MP3 Format

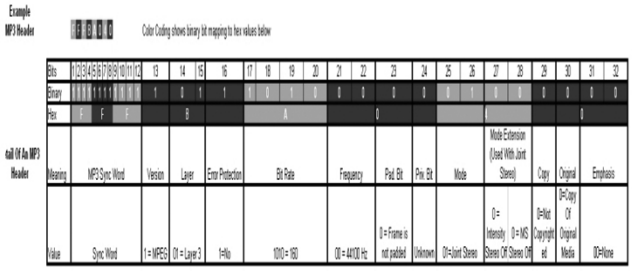

Figure 4: Example of MP3 Header

The comments section of the ID3 tag of the current MP3 format can contain equalization settings for various song lengths. MP3 players can be configured to decode these settings and adjust them accordingly while the track is playing.

III. SUGGESTED APPROACH:

To improve sound quality, and to limit the physical extremes that would arise from recording analogue records, equalization of audio is very helpful. Most of the MP3 players have an embedded graphic equalizer which is a high-fidelity audio control that allows the user to see graphically and control individually a number of different frequency bands. The equalizer settings have been maintained in semicolon separated fields for varying time durations; for a certain time duration, the gainlevel values have been recorded in colon separated fields. The equalization settings are kept in the comments field of the MP3 format's id3 tag. The new format can be played on the current MP3 players without changing the MP3 Header MP3 Data MP3 Header MP3 Data MP3 Header MP3 Data ID3v2x Metadata Existing MP3 Format Equalizer Settings Embedded here New MP3 Format equalizer, and the old MP3 format can also be played on the new MP3 player with the equalizer set to default values because care has been taken to preserve the audio information contained in the existing MP3 format. A program has been created that can convert MP3 files currently in existence into the new format, recover equalizer settings from any file in the new format, and play newly created MP3 files in a player by properly adjusting the equalizer while the music is playing [6,7].

There are four stages to the implementation portion. Phase One: An application was created to insert equalizer settings within the pre-existing MP3 file, converting it from one format to another. Phase Two: The MP3 player was configured to correctly decode and configure these parameters. This also entails giving Windows Media Player a fresh look. Third Phase: The player was reprogrammed to decode the time-varying equalizer settings and set the equalizer properly. Fourth Phase: The application was enhanced and the front end was improved.

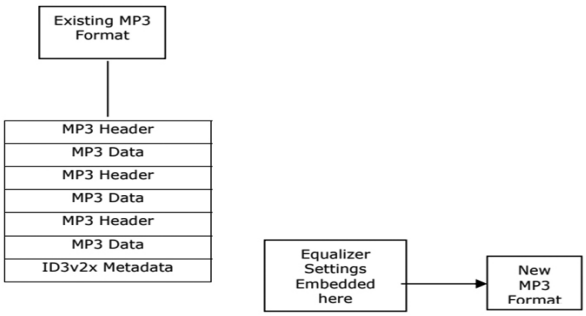

Figure 5: conversion of Existing MP3 Format to new MP3 Format

The implementation part has been divided into four phases. First Phase: An application was designed to convert the existing MP3 into the new format i.e. equalizer settings were embedded into the already existing MP3 format. Second Phase: The MP3 player was programmed to decode and set these settings appropriately. This also includes creating a new skin for windows media player. Third Phase: The equalizer settings were made time-varying and the player was reprogrammed to decode these settings and set the equalizer appropriately. Fourth Phase: Front end was improved and the application was enhanced.

After that, the modules were combined, and the finished program was tested and debugged. Phase I The project's first phase has been implemented using VB.Net and CDDBControl.dll, which provides an interface to the ID3 tag of the currently used MP3 format. Inside CDDBControl.dll is the CddbID3Tag class, which has the following data members: BeatsPerMinute as String, Album as String.

This string contains the following: Comments, String CopyrightHolder, String CopyrightYear, FileId, Genre, ISRC, Label, Lead Artist, Movie, PartOfSet, TrackPosition, Year, String.

A ten-band equalizer with the following frequency bands has its equalizer settings stored in the comments section of the ID3 tag: 31 Hz, 62 Hz, 125 Hz, 250 Hz, 500 Hz, 1 Hz, 2 Hz, 4 Hz, 8 Hz, and 16 Hz. For each band, the gain level has been stored in the comments section of the ID3 tag for varying time durations; for a given time duration, the values for gainlevels have been stored in colon separated fields. In order to do this, the source file is copied to the target URL and then formatted to the new one.

Additionally, this application uses VB.NET programming to build an aesthetically pleasing arrangement of subjective phrases for the equalization controls, making them user-friendly even for inexperienced users. The application offers the ability to scroll through, pause, stop, and play MP3 songs in order to determine the proper equalizer settings. The Windows Media Player ActiveX component has been used to do this.

3.2.2. Phase II: A new skin for Windows Media Player has been created in order to implement phase II. Skins can be used to add functionality to Windows Media Player that is not already there. The Windows Media Player 9 Series SDK's primary interop assembly, Wmppia.dll, is made available by Microsoft and makes up the WMPEqualizerSettingsCtrl class. The equalizer has been adjusted to the gain levels kept in the new mp3 format using the EqualizerSettings element. An event is triggered whenever a media file is opened via the skin, and Jscript handles this event. Before playing the media, this event handler changes the equalization to the settings included in the new mp3 format. The following files make up the skin: • Skin Definition File, a master file with the.wms file extension that specifies how the other files are to be used. The basic instructions for the functions of the skin and the locations of other files it uses are contained in the skin definition files. Extension Markup Language (XML) is used in the skin definition file to write the instructions. • Art Files: These files hold the skin's graphic components. BMP, GIF, JPEG, and PNG formats are among them.

JScript Files: Script files, which have the.js file extension and are generated with Microsoft Jscript, are used to construct more intricate reactions to events. The following photos are included in the art files: • Primary photos: Users view these images after installing the skin. One or more images produced by particular skin controls make up the main image. • Mapping Images: These pictures are employed in image mapping to set off skin-related actions. However, Windows Media Player uses the graphics in an image map file to trigger actions when the user clicks on the skin; they are not intended for user viewing. • Alternate graphics: Another option is to program other graphics to appear in response to user actions. For instance, while the mouse is over a button, an alternate image of the button will appear. 3.2.3 Phase III: The player was reprogrammed to decode the time-varying equalizer settings and adjust the equalizer accordingly. The time-varying equalizer settings are recorded in semicolon separated fields for different time durations, and the values for gainlevels have been stored in colon separated fields for a certain time duration. The comments field of the id3 tag of the MP3 format has been utilized to store these settings. With the aid of the VB.NET CDDBControl.dll, settings are integrated into the current MP3 format.

The equalization settings for various time durations are extracted first, and the value for each gainlevel is then extracted from these settings, which is how the player has been reprogrammed to decode these settings. The WMS file uses a timer,

and the Jscript file handles the ontimer event. It compares the current track position to the time period of the settings retrieved, and if a match is discovered, sets the equalization to the extracted settings for that specific duration.

3.2.4 Phase IV: During this stage of the project, the application's front end was made more user-friendly, several menu-driven controls were added, and a comprehensive project help file was produced. Additionally, there was a way to access and retrieve the equalization configuration that was saved in any MP3 file in the updated format. Additionally, the ability to construct playlists in HMTL, TXT, and M3U formats was introduced. The new player, which was created in phase three, will also support this playlist. As a result, it is possible to create a playlist that contains both the new and the old MP3 files. When this playlist is chosen for playback, the equalization for the files in the older MP3 format is set to the factory defaults, and for the files in the newer MP3 format, the settings are determined by the parameters inherent in the format.

CONCLUSION

In order to save consumers from having to manually adjust the equalizer every time they listen to the same MP3 music, a new MP3 format has been proposed that may include embedded equalizer settings for various track durations. The MP3 player is designed to decode the contained equalization settings and adjust them appropriately while a track is playing. The equalizer settings are part of the MP3 format tags. The new format may be played on existing MP3 players without the equalization being set, and the old MP3 format can also be played on the new MP3 player with the equalizer set to default settings because care has been made to preserve the audio information included in the old MP3 format. A program has been created that can convert MP3 files currently in existence into the new format, get equalizer settings from any file in the new format, and play newly created MP3 files ina player by automatically adjusting the equalizer while the music is playing. Additionally, playback and playlist creation capabilities has been included. To play files in the new MP3 format, a new skin for Windows Media Player has been created, which automatically changes the equalizer to settings encoded in the new MP3 format.

V. FUTURE SCOPE

In addition to equalization, there may be other improvements that might be made to the current MP3 standard. In a similar way, this work can also be expanded to support other audio formats, such WAV, WMA, etc.

REFERENCES

[1] A study on the influence of senses and the efficacy of sensory branding was conducted by Rupini RV and Nandagopal R in the Journal of Psychiatry in 2015. DOI: 10.4172/1994-8220.1000236

[2] Deep Learning Based Music Recommendation System, Malige Gangappa, Avuluri Nikhitha, Bondugula Reshma, Gundabathina Manasa, Pooja Kumari Singh, International Research Journal of Engineering and Technology (IRJET), Volume: 10 Issue: 02 | Feb 2023

[3] Xin H. "Enhanced Deep Neural Network-Based Attention Mechanism-Based Music Recommendation Algorithm," Hindawi Mobile Information Systems, Volume 2022, Article ID 4112575, 11 pages (https://doi.org/10.1155/2022/4112575).

[4] Yunzhe Dong, "A Machine Learning-Based Music Recommendation System," AMMMP 2023, Volume 47 (2023)

[5] The article "Music Recommendation Using Deep Learning A Study" was published in the International Journal of Creative Research and Thoughts in May 2022 (Volume 10, Issue 5), by Shubham Kedari, Jui Walimbe, Sakshi Warule, and Madhuri Thorat.

[6] Markus Schedl, Frontier in Applied Mathematics and Statistics, 2019, "Deep Learning in Music Recommendation System."

[7] "Feature-combination hybrid recommender systems for automated music playlist continuation," Andreu Vall, Matthias Dorfer, Hamid Eghbal-zadeh Markus Schedl, Keki Burjorjee, and Gerhard Widmer, User Modeling and User-Adapted Interaction (2019) 29:527–572, https://doi.org/10.1007/s11257-018-9215-8

Assistant Professor **Kamna Singh** works at the Ajay Kumar Garg Engineering College in Ghaziabad. Her degree is B.Tech. from JIIT, Noida, and MTech from UCER, Allahabd. Internet of Things, cyber security, information security, computer networks, and data mining are among the topics she studies. She has numerous information security publications published in reputable journals. She has multiple research publications published in international journals with Scopus indexing.