SPEECH UNDERSTANDING
(CSL7770)

Comparative Spectrographic Analysis of Songs Across Four Different Genres

ASSIGNMENT TASK-B REPORT
(GitHub Link)

SUBMITTED BY

ARNAV SHARMA

(M24CSE004)

Department of Computer and Engineering

Indian Institute of Technology Jodhpur

Karwar 342030, Rajasthan, India

ABSTRACT

This task analyses the spectrograms of four different songs, each representing a distinct musical genre: EDM, Hip-Hop, Pop, and Rock. The objective is to understand the spectral characteristics of these genres by visualizing their frequency distribution over time. The analysis includes comparisons based on spectrogram shape, frequency distribution, power levels, and general sound characteristics. The results show distinct differences in spectral energy distribution and intensity, reflecting the unique production styles and instrumentation of each genre.

INTRODUCTION

Music genres exhibit unique spectral properties due to differences in instrumentation, vocal processing, and production techniques. This study focuses on four songs from four genres: "Bangarang" by Skrillex representing EDM, "Lose Yourself" by Eminem representing Hip-Hop, "Perfect" by Ed Sheeran representing Pop, and "Thunderstruck" by AC/DC representing Rock. A spectrogram visually represents the frequency content of an audio signal over time, making it a valuable tool for understanding these differences. By analysing the spectrograms of these songs, it is possible to gain insights into the structure and composition of each genre. The study aims to highlight how frequency content, energy levels, and temporal characteristics vary among these musical styles.

METHODOLOGY

The first step in the analysis involved collecting audio data from a local directory and selecting songs that best represent their respective genres. Each audio file was then preprocessed to ensure uniformity, including resampling all signals to a standard 16 kHz sample rate and retaining only the first audio channel for consistency. The spectrograms were computed using the Short-Time Fourier Transform (STFT) with an FFT size of 1024 and a hop length of 512. The transformation of spectrograms into a logarithmic scale (dB) enhanced visualization, making it easier to compare differences in power distribution. Once the spectrograms were generated, they were analyzed for their statistical properties, such as minimum and maximum values and mean power levels. These properties provided deeper insights into the spectral energy distribution of each song.

RESULTS AND DISCUSSIONS

The analysis of the EDM track "Bangarang" by Skrillex revealed a high spectral density with strong energy concentrated in the upper-frequency range. The spectrogram showed noticeable breaks corresponding to beat drops, which are characteristic of the EDM genre. The maximum value observed in the spectrogram was 30,528, and the mean power was measured at -2.44 dB, indicating significant loudness and energy. In contrast, "Lose Yourself" by Eminem, representing the Hip-Hop genre, exhibited a lower spectral density but maintained a broad frequency range. The mid-frequency content was dominant, reflecting the prominence of

speech in Hip-Hop music. Despite having the highest maximum value of 34,942, the mean power was lower at -10.59 dB, suggesting more dynamic variation in energy levels.
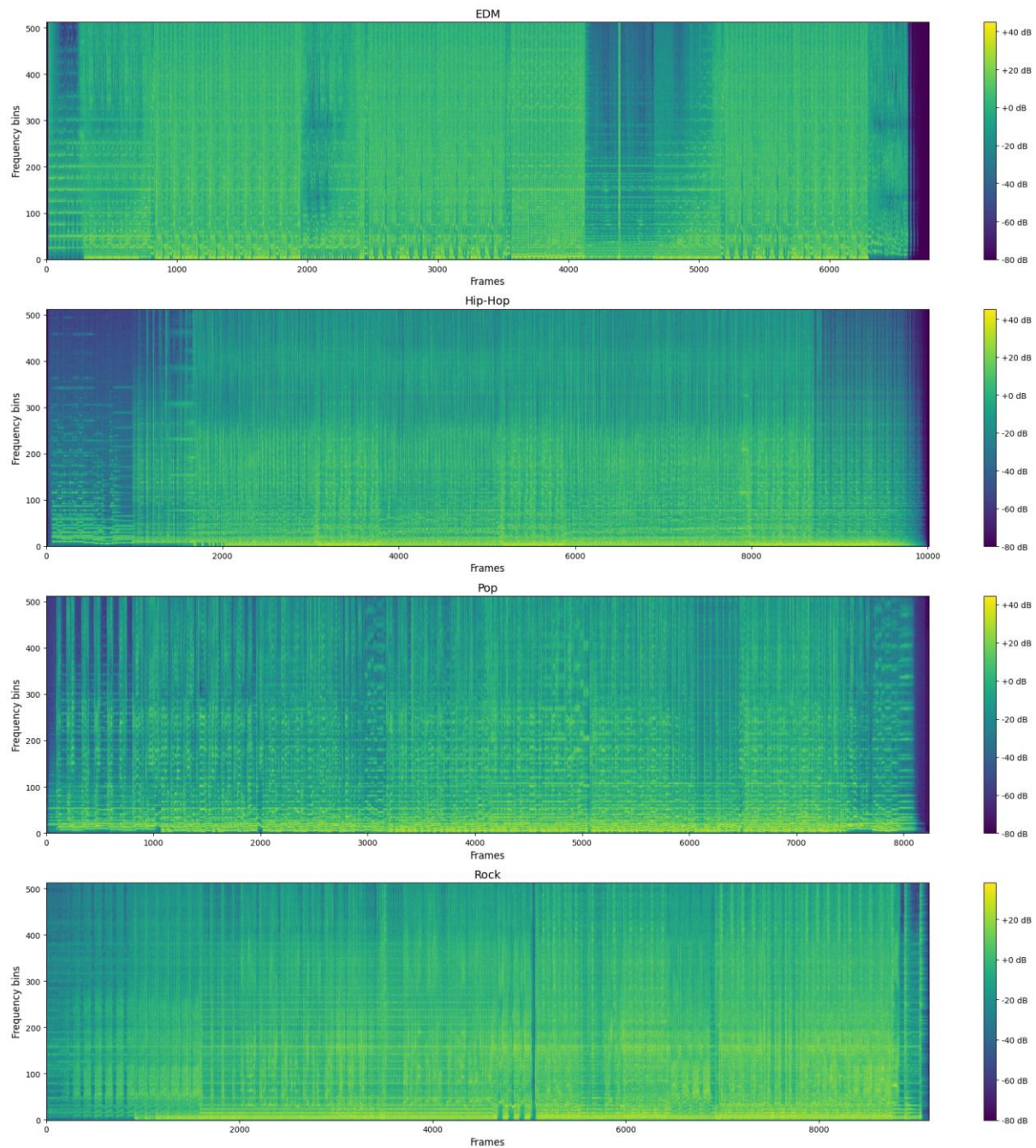

Figure 1 Spectrograms of different song genre

The Pop song "Perfect" by Ed Sheeran displayed a more evenly distributed spectral content, particularly across mid and low frequencies. The structure appeared smoother with fewer sharp transitions, indicating a more melodic and harmonic composition. The maximum value observed was 28,713, with a mean power of -9.42 dB, reflecting a moderate loudness level with a balanced sound. On the other hand, "Thunderstruck" by AC/DC, representing the Rock genre, showed a strong presence in mid and high frequencies, likely due to the dominance of electric guitars. Unlike EDM, which featured pronounced energy bursts, the Rock spectrogram

demonstrated a more sustained energy level throughout the song. The maximum value of 6,511 was the lowest among the analysed tracks, but the mean power was relatively high at -4.06 dB, highlighting the continuous intensity typical of Rock music.

When comparing the spectrograms of these four genres, clear distinctions emerged. EDM and Hip-Hop displayed stronger transient energy variations, while Pop and Rock exhibited smoother transitions. Hip-Hop had the most dynamic spectral distribution, with noticeable fluctuations between high and low energy levels. In contrast, Rock maintained a consistent energy level throughout the track, aligning with its genre characteristics. These findings emphasize the role of production techniques in shaping the spectral properties of different musical styles.

CONCLUSIONS

The spectrogram analysis conducted in this study illustrates the distinct spectral properties of different music genres. EDM is characterized by high-energy bursts and wide frequency coverage, while Hip-Hop emphasizes rhythmic and speech-like structures. Pop maintains a smooth and balanced frequency spectrum, making it ideal for melodic compositions, whereas Rock features strong mid-frequency content with sustained intensity due to electric guitars and percussion. These findings highlight the importance of spectrogram analysis in music genre classification and audio signal processing. Future research could expand this study by analysing a larger dataset and incorporating machine learning techniques for automatic genre recognition, further enhancing the understanding of how different musical styles are constructed and perceived.