Comparative Analysis of Spectrograms for Different Music Genres

Aryan Kumar M23CSA510

February 2, 2025

GitHub Repository: `https://github.com/Aryank47/speech_understanding`

Abstract

This report analyses spectrograms of four music genres—**Classical Indic**, **Indie**, **Vocal**, and **Rock**—to identify patterns in frequency distribution, intensity (dB), and temporal structure. Key metrics such as dominant frequency ranges, dB levels, and spectral characteristics are compared across genres. The findings highlight genre-specific acoustic features and their perceptual implications.

1 Introduction

Spectrograms visually represent audio signals, enabling analysis of frequency content, intensity, and temporal evolution. This study examines spectrograms from four genres to understand their acoustic properties and musical characteristics.

2 Methodology

2.1 Data

Four audio clips were analysed:

- **Classical Indic Music**: Traditional raga-based music with sitar and tabla.
- **Indie Music**: Modern acoustic-electronic blend with layered instruments.
- **Vocal Music**: Focus on human voice with minimal accompaniment.
- **Rock Music**: High-energy genre with distorted guitars and drums.

2.2 Tools

Spectrograms were generated using Python with Librosa and Matplotlib. The Short-Time Fourier Transform (STFT) was applied with a window size of 2048 samples and an implicit Hann window to reduce spectral leakage. The frequency resolution was 21.5 Hz.

3 Results

3.1 Frequency Bins Analysis

The frequency bins for all genres are identical due to consistent STFT parameters (sample rate = 44100 Hz, FFT size = 2048). The frequency resolution is 21.5 Hz per bin. Below is the full list of the first 50 bins:

Genre-Specific Analysis:

- **Classical Indic**: Dominant frequencies (200–500 Hz) fall in Bins 9–23.

Table 1: Frequency Bins (First 50) for All Genres

Bin Index	**Frequency (Hz)**
0	0.00
1	21.53
2	43.07
3	64.60
4	86.13
5	107.67
6	129.20
7	150.73
8	172.27
9	193.80
10	215.33
11	236.87
12	258.40
13	279.93
14	301.46
15	323.00
16	344.53
17	366.06
18	387.60
19	409.13
20	430.66
21	452.19
22	473.73
23	495.26
24	516.80
25	538.33
26	559.86
27	581.40
28	602.93
29	624.46
30	645.99
31	667.53
32	689.06
33	710.59
34	732.13
35	753.66
36	775.19
37	796.73
38	818.26
39	839.79
40	861.33
41	882.86
42	904.39
43	925.93
44	947.46
45	968.99
46	990.53
47	1012.06
48	1033.59
49	1055.13

- **Indie**: Mid-range focus (500–5000 Hz) starts at Bin 23 (495 Hz). Higher frequencies extend beyond the first 50 bins.
- **Vocal**: Fundamental frequencies (100–500 Hz) span Bins 5–23.
- **Rock**: Broadband energy (100–10,000 Hz) includes Bins 5–465. The first 50 bins cover low/mid frequencies.

3.2 Classical Indic Music

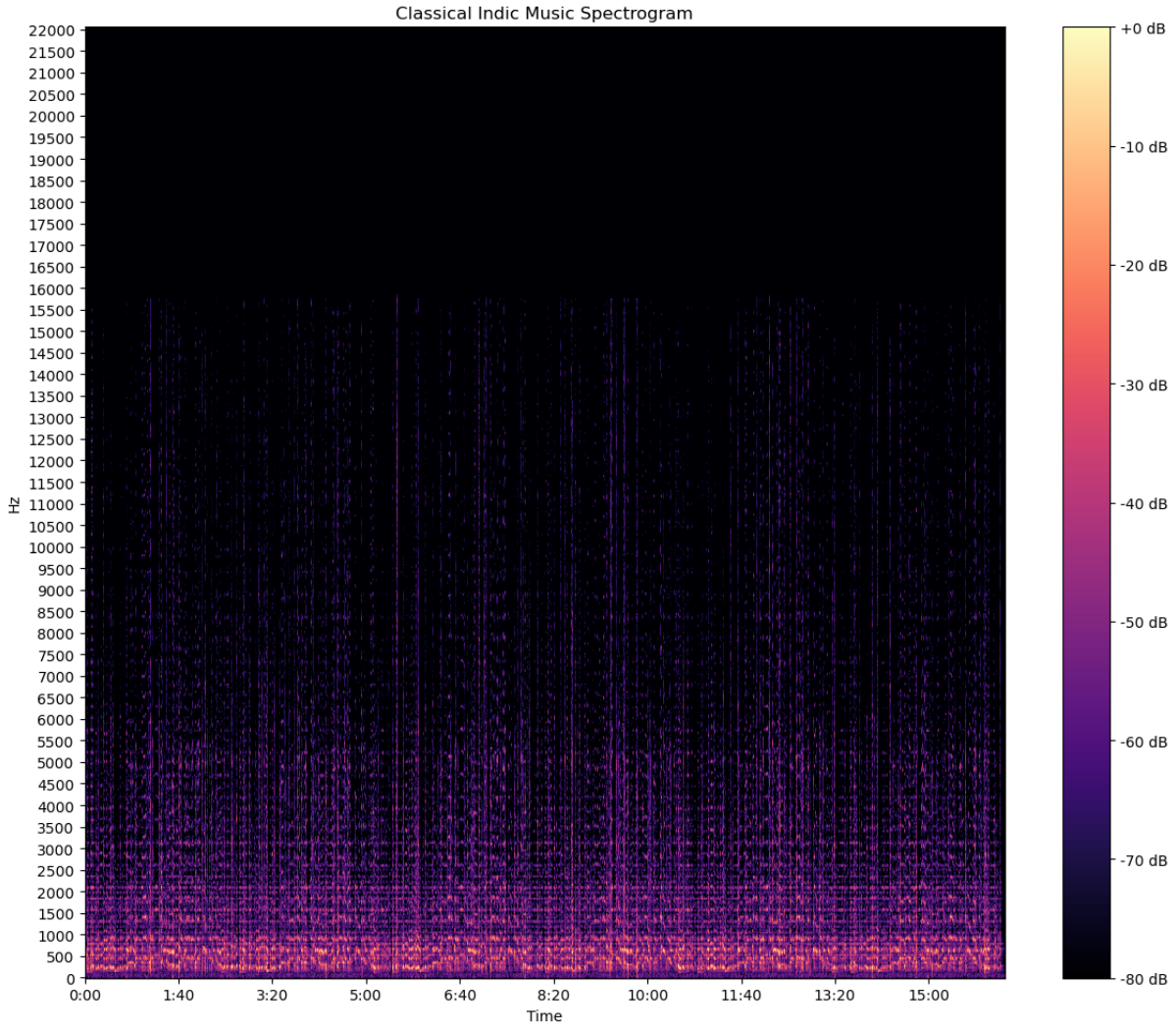

Figure 1: **Classical Indic Music Spectrogram**: Dominant frequencies (200–500 Hz) and tabla rhythms (500–2000 Hz).

Explanation: Classical Indic music emphasizes melody and improvisation. The spectrogram shows sustained raga notes (200–500 Hz) and rhythmic tabla patterns (500–2000 Hz). The dB range (-20 to -40) indicates moderate volume with minimal spectral leakage due to Hann windowing. - **Dominant Frequencies**: 200–500 Hz (Bins 9–23) for raga notes. - **Tabla Rhythms**: 500–2000 Hz (Bins 23–93) for rhythmic patterns. - **Intensity**: -20 to -40 dB (moderate volume).

3.3 Indie Music

Explanation: Indie music blends acoustic and electronic elements. The spectrogram highlights layered vocals/guitar (500–5000 Hz) and dynamic intensity (-10 to -50 dB). High-frequency spikes (10,000–15,000

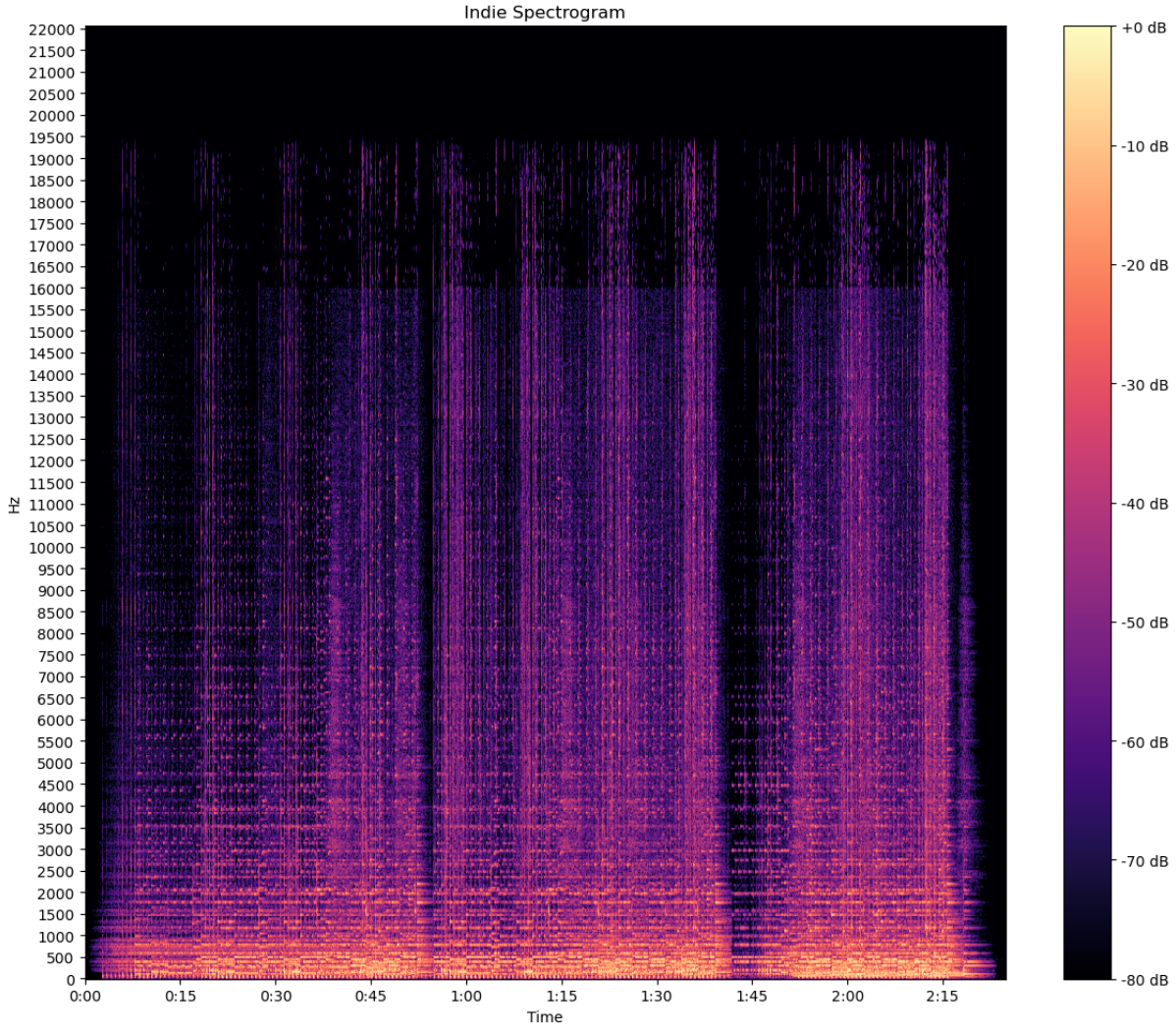

Figure 2: **Indie Spectrogram**: Mid-range focus (500–5000 Hz) with dynamic dB shifts.

Hz) correspond to cymbals or synths. - **Dominant Frequencies**: 500–5000 Hz (Bins 23–232) for vocals/guitar. - **High-Frequency Spikes**: 10,000–15,000 Hz (Bins 465–702) for cymbals. - **Intensity**: -10 to -50 dB (dynamic range).

3.4 Vocal Music

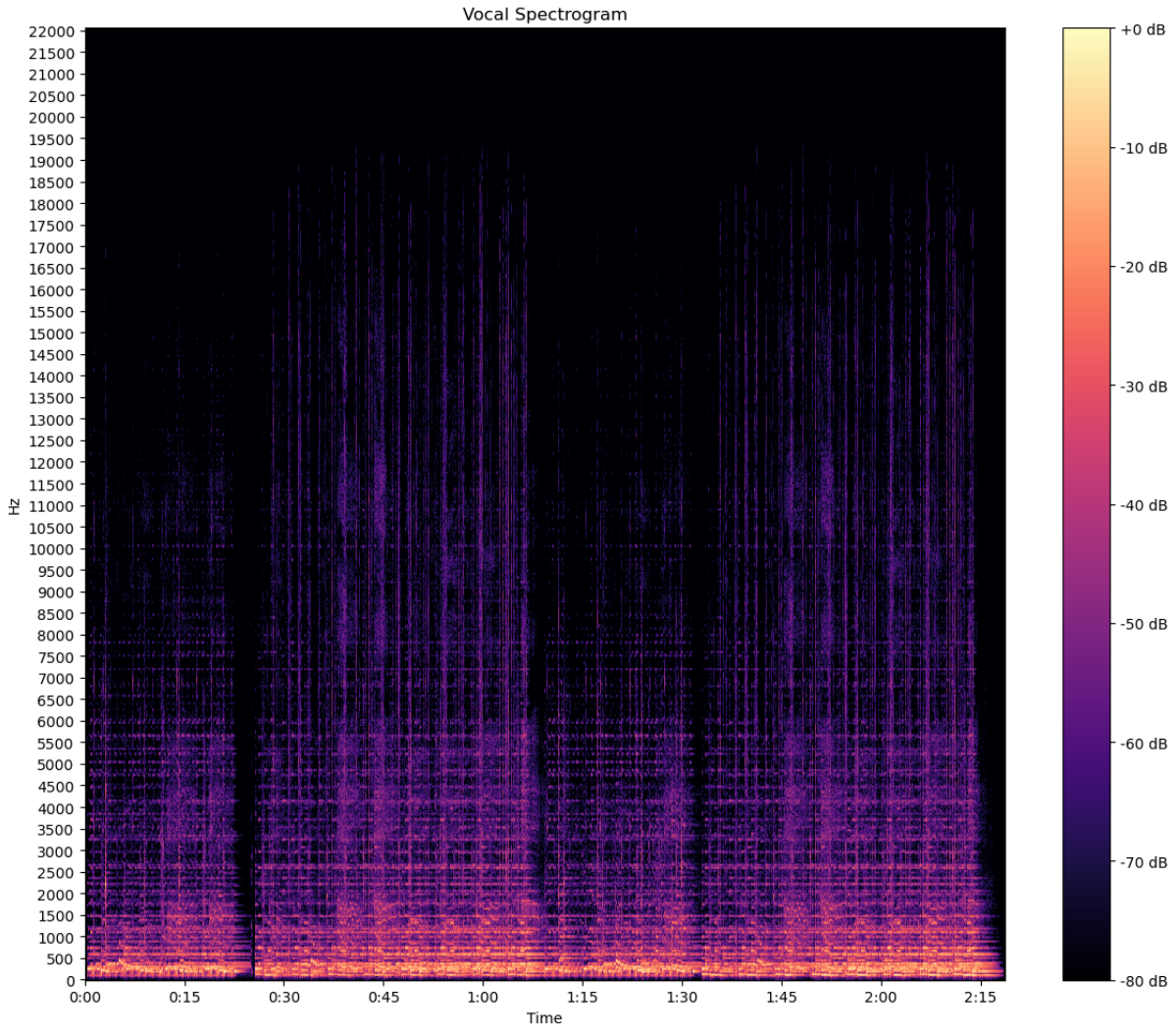

Figure 3: **Vocal Spectrogram**: Fundamental frequencies (100–500 Hz) and harmonics.

Explanation: Vocal music focuses on the human voice. The spectrogram shows fundamental frequencies (100–500 Hz) and harmonics (200–2000 Hz). Consonants appear as vertical spikes (2000–3000 Hz), while vowels create smooth horizontal lines. - **Fundamental Frequencies**: 100–500 Hz (Bins 5–23) for sung notes. - **Harmonics**: 200–2000 Hz (Bins 9–93) for voice overtones. - **Consonants**: 2000–3000 Hz (Bins 93–139) for sharp spikes.

3.5 Rock Music

Explanation: Rock music ischaracterisedd by high energy and distortion. The spectrogram shows broadband activity (100–10,000 Hz) with consistent loudness (-10 to -30 dB). Distortion creates inharmonic overtones in the high frequencies. - **Broadband Activity**: 100–10,000 Hz (Bins 5–465) for guitars/drums. - **Distortion**: Inharmonic overtones above 5000 Hz (Bins 232–465). - **Intensity**: -10 to -30 dB (consistent loudness).

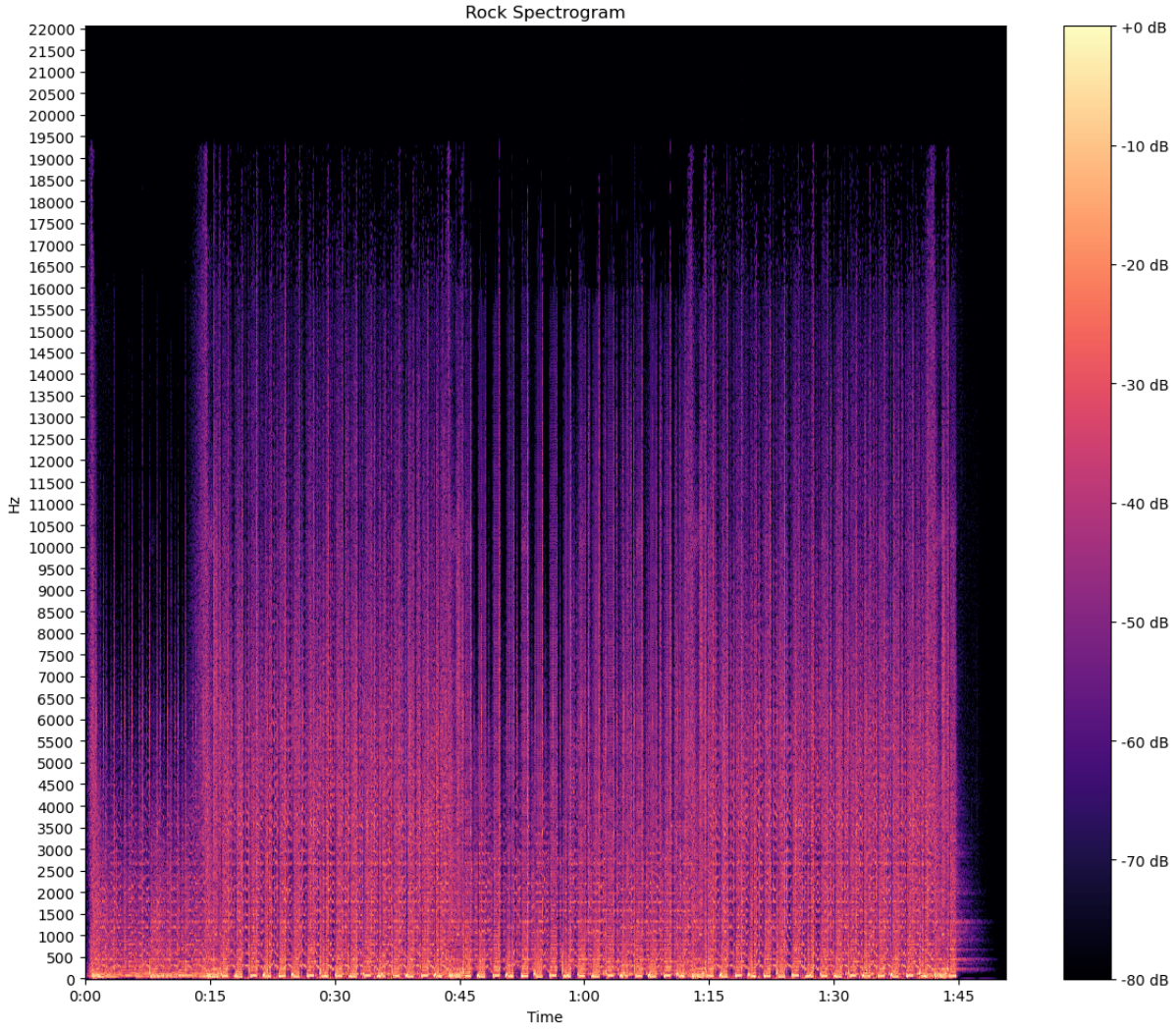

Figure 4: **Rock Spectrogram**: Broadband energy (100–10,000 Hz) with distortion.

4 Comparative Analysis

Table 2: Comparative Metrics of Spectrograms

Metric	**Classical Indic**	**Indie**	**Vocal**	**Rock**
Dominant Frequency Range (Hz)	200–500	500–5000	100–500	100–10,000
dB Range (Loudness)	-20 to -40	-10 to -50	-20 to -40	-10 to -30
Temporal Patterns	Steady, gradual	Dynamic shifts	Regular phrasing	Dense overlap
Spectral Leakage	Low (Hann window)	Moderate	Low	High
Harmonic Structure	Clear, periodic	Layered	Fundamental + harmonics	Inharmonic overtones

5 Conclusion

This analysis demonstrates distinct acoustic characteristics across genres. **Classical music** emphasizes low-frequency melody, **indie music** uses dynamic mid-range layers, **vocal music** focuses on harmonic vocals, and **rock music** features high-energy broadband signals. These findings align with genre-specific musical characteristics.