WAVELET-BASED MUSIC GENRE CLASSIFICATION

Group - 10

Presented By

PALANIAPPAN AN (CB.EN.U4CCE21042)

PRANAV J (CB.EN.U4CCE21046)

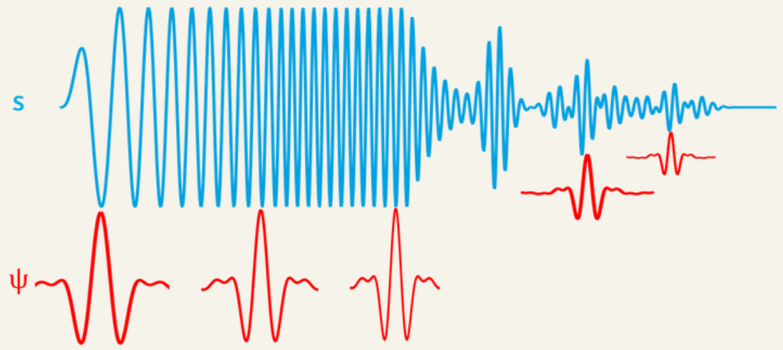
VISHVESHWARAN (CB.EN.U4CCE21078)

Audio Coding

Audio coding using wavelets involves compressing audio signals by transforming them into a more efficient representation while retaining perceptual quality. Wavelet transforms play a crucial role by offering multi-resolution analysis, which captures both time and frequency characteristics of the audio signal.

How does wavelet analyze frequencies in audio?

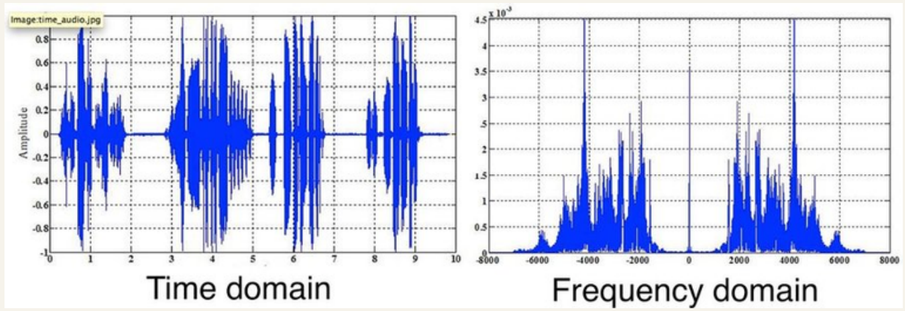
- Time-Frequency Localization
- Wavelet Transform Process
- Filter Bank Technique
- Frequency Bands Representation



Representation using different domains

1. **Time Domain**

- The raw audio signal as a function of amplitude over time**.**
- Useful for analyzing the signal's overall shape and amplitude envelope.
- Insufficient for detailed frequency analysis.



2. **Frequeny Domain**

- Obtained via Fourier Transform, showing how the signal's energy is distributed across frequencies
- Reveals harmonic content and spectral properties.
- Loses time information, making it hard to detect transient events.

3. **Time-Frequency Domain (Wavelet Domain)**

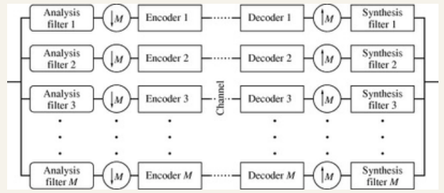
- Combines time and frequency representations, providing a localized view of which frequencies are present at specific times.
- High-resolution analysis of both transient and stationary components.
- Example: Detecting when a drumbeat (high-frequency) occurs or tracking bassline patterns (low-frequency).

4. **Scalogram**

- A visual representation of the wavelet transform's output, plotting scale (frequency) versus time with amplitude as intensity.
- Highlights transient features and frequency shifts over time.

Popular methods used in audio coding

- **Subband Coding (SBC)**
  - Uses filter banks to split the audio signal into frequency bands.
  - Compresses each band separately using techniques like quantization.
  - Example: MP3 coding.

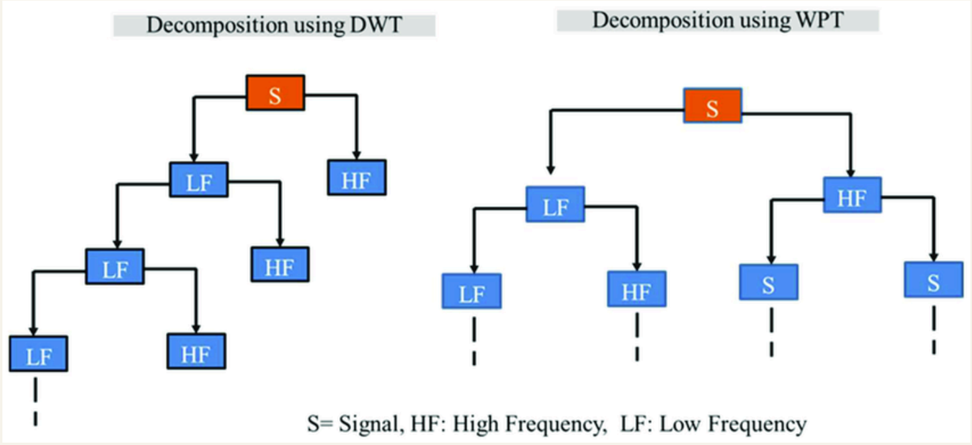


- **Transform Coding**
  - Uses transforms (like DWT or DCT) to represent audio in the frequency domain, exploiting redundancy.
  - Coefficients are quantized and encoded.
  - Wavelet transform allows multi-resolution representation.

- **Embedded Zero-tree Wavelet (EZW) Coding**
  - Compresses wavelet coefficients by representing large coefficients and their hierarchical relationships efficiently.
  - Suitable for progressive transmission.

- **Set Partitioning in Hierarchical Trees (SPIHT)**
  - Enhances EZW by using a tree-based structure to encode significant coefficients.
  - Provides high compression ratios with scalability.

- **Wavelet Packet Transform (WPT)**
  - An extension of the DWT that provides finer frequency resolution.
  - Encodes audio signals more precisely by adaptively decomposing the signal based on its characteristics.



Objective

To classify music genres by analyzing audio signals through Discrete Wavelet Transform (DWT), extracting key features from the time-frequency domain to identify genre-specific patterns.

Dataset

- The GTZAN dataset contains 1,000 audio files across 10 genres, with each track being 30 seconds long, widely used in music genre recognition research.

- The dataset includes Mel Spectrogram images and two CSV files with extracted audio features: one with mean and variance over 30-second tracks and another with features from 3-second segments for increased data volume.

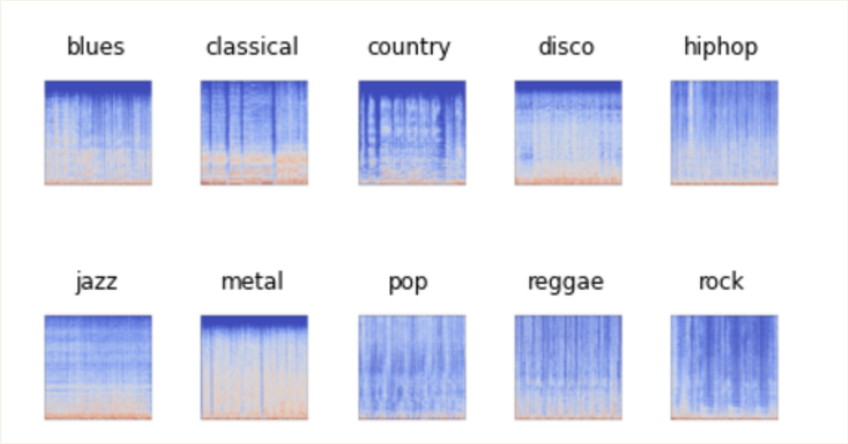
Methodology

1. Data Collection & Preprocessing

2. Representation using different domains

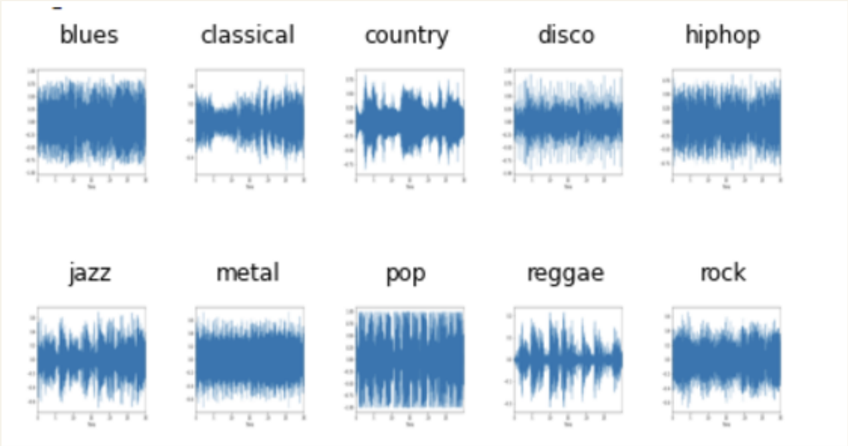
3. Wavelet Decomposition using DWT

4. Plots of different Genre



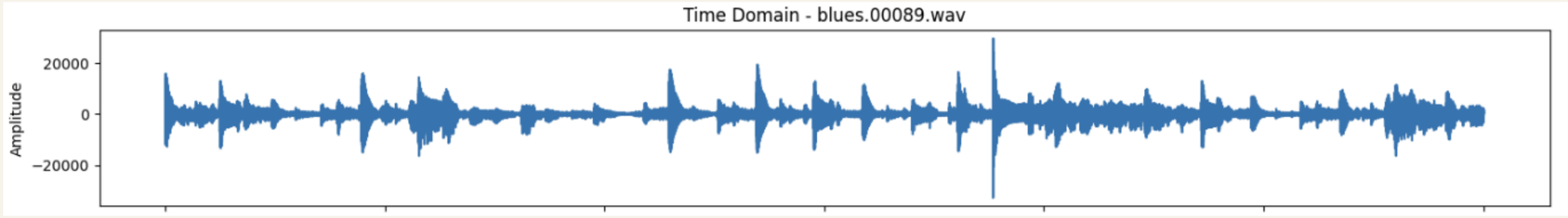
Libraries used

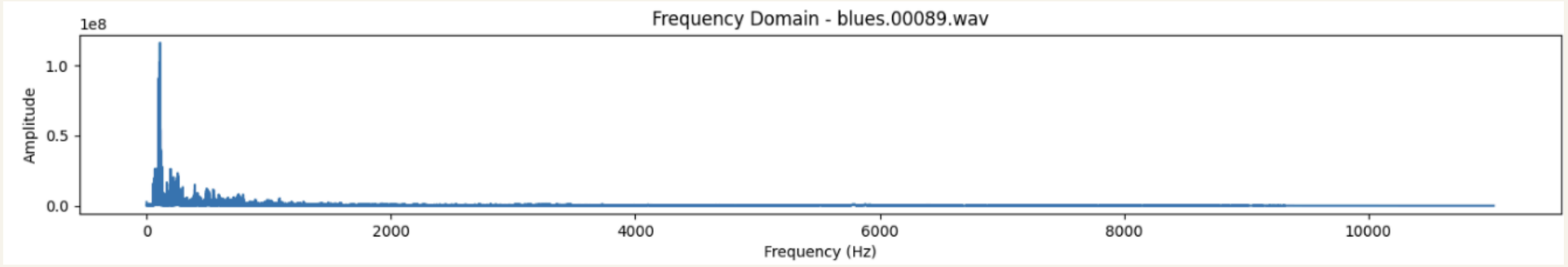
- Scipy
- pywt
- numpy
- matplotlib

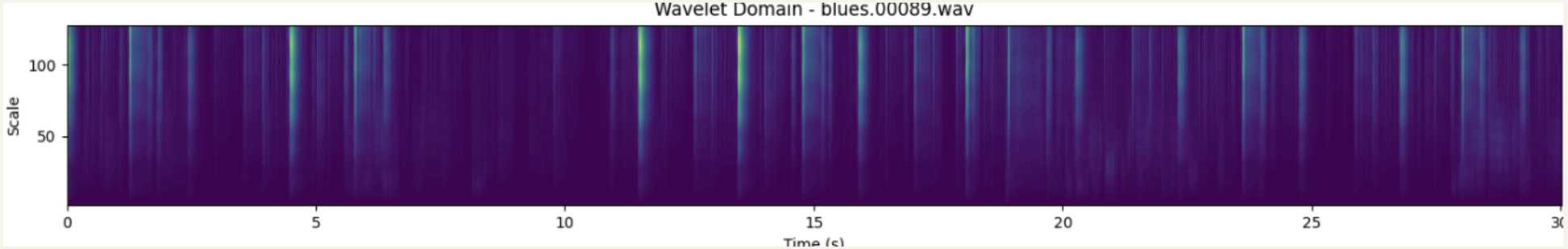


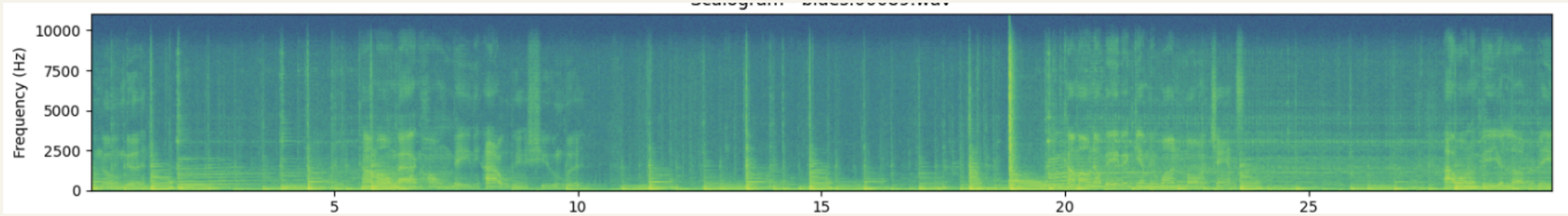
Representation

Blues

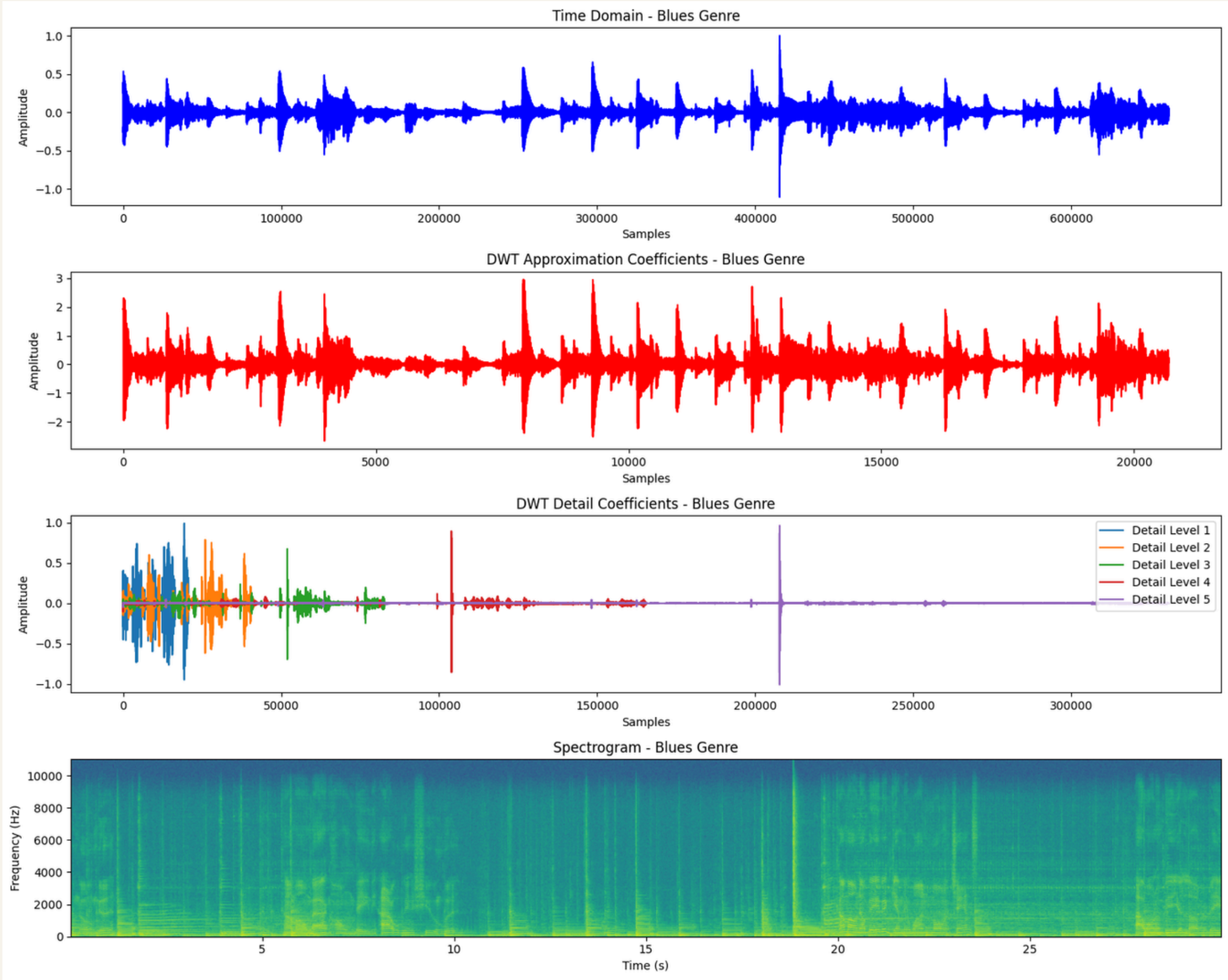








Discrete Wavelet Transform



Characteristics of Blues

- Harmonic Structure: Repetitive 12-bar blues progression with predictable chord changes that create emotional resonance.
- Melodic Expression: Emphasis on blue notes (flattened 3rd, 5th, 7th), adding a soulful and melancholic tone.
- Rhythmic Features: Swing or shuffle rhythms paired with call-and-response patterns between vocals and instruments.

THANK YOU