Executive Summary: The Embedder's New Groove

Duncan Wood, Harrison Chen, Utkarsh Agrawal

December 6, 2024

Our project focuses on developing a method to encode and analyze the rhythmic content of musical recordings. The primary objective is to create an embedding that captures the nuances of musical grooves—patterns of rhythm and timing—using classical signal processing and clustering techniques. This encoding enables objective comparisons across and within musical genres, in hopes of discovering new musical insights, and applications in music recommendation engines.

Objectives

- **Rhythm Classification:** Use unsupervised methods to identify and group distinct rhythmic patterns (e.g., samba, swing, EDM).
- **Encoding Grooves:** Develop a representation that encapsulates timing, emphasis, and other subtle rhythmic features.
- **Generative Applications:** Extend the embedding to enable groove transformation, such as altering a recording's style.

Potential Applications

- Analyzing rhythmic trends in popular music over time.
- Auto-DJ systems for rhythmically similar track selection.
- Tools for music producers to generate and manipulate grooves.
- Applications in copyright detection and music recommendation systems.

Stakeholders

This work is relevant to the music industry, including producers, content creators, streaming services, and event planners.

Methodology

1. **Downbeat Detection:** Tools like BeatNet were used to identify beat locations, which segment the music into measures. However, poor accuracy ($\sim$50%) limits scalability.
2. **Groove Embedding:**
    - Divide measures into frequency bands (low, mid, high).
    - Process signal power data and integrate over time subdivisions.
    - Represent each measure with a high-dimensional vector reduced via PCA for clustering.
3. **Clustering:** Employ Gaussian mixture models to identify representative rhythms within and across recordings.

Results

- The embedding captures within-song rhythmic variations and representative measures.
- Initial tests demonstrated clustering potential within individual songs but struggled to distinguish broader genres due to inaccuracies in beat detection.

Challenges and Next Steps

1. **Improving Beat Detection:** Develop improved techniques or leverage better-labeled datasets to refine downbeat accuracy.
2. **Instrument Differentiation:** Enhance the encoding to differentiate between rhythmic components (e.g., bass drum, snare, hi hats).
3. **Microtiming Analysis:** Implement methods to analyze small timing variations relative to a perfect rhythmic grid, a key feature for understanding stylistic nuances in grooves.
4. **Genre-Level Insights:** Train algorithms to recognize families of grooves across genres, addressing the limitations of Euclidean clustering.