# Music Genre Classification
Alex Stabile

Abstract

The aim of this project is to enable a computer to place given pieces of music into their appropriate genres. Genres are often generalities that do not necessarily fit a piece exactly, but this project could determine what music really is similar and should be grouped together. This could have applications in sorting large libraries of music or suggesting music to individuals based on what they like to listen to. Music in midi format is used as input for the program. Python was used to write classes that can read and store the information held by midi files. This is done by grouping notes into their appropriate beats. Organizing them in such a manner allows for harmonic analysis, which is a good indicator of genre. Data about harmonies and other musical characteristics will be analyzed by a statistical model to determine what elements different genres have in common.

Background

Current research often uses statistical models to determine how a given piece of music should be categorized. One approach attempted an automatic method of classification that is completely general (Cilibrasi, Vitanyi, Wolf). This was done by looking for mathematical similarity, rather than features specific to music. Midi files were used for musical analysis, and very successful results were shown when grouping music into rock, jazz, or classical genres. Results were moderately successful when attempting to group pieces by composer, but got worse as sample size increased. Interestingly, the algorithm could even cluster like file types together (sorting out java class files, gene sequences from different species, and widely different styles of music).

No method has yielded perfectly successful results – even many people cannot successfully place music (especially from a short sample) into its correct genre.

One especially successful experiment was "On musical stylometry—a pattern recognition approach" (Backer et al). Different musical aspects of pieces were analyzed, and a statistical model was created to group new pieces into their appropriate period. By analyzing more musical characteristics, their model became more fine-tuned. These characteristics included harmonies, dissonance, note entropy, and types of intervals.
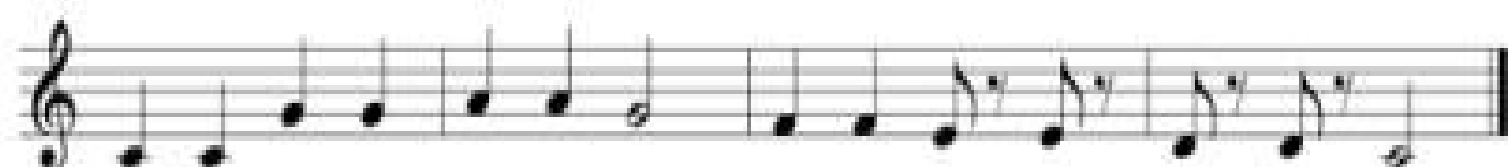

(Delta time, MIDI protocol bytes)

```
00 ff 58 04 04 02 30 08 00 ff 59 02 00 00 00 90 3c 28
81 00 90 3c 00 00 90 3c 1e 81 00 90 3c 00 00 90 43 2d
81 00 90 43 00 00 90 43 32 81 00 90 43 00 00 90 45 2d
81 00 90 45 00 00 90 45 32 81 00 90 45 00 00 90 43 23
82 00 90 43 00 00 90 41 32 81 00 90 41 00 00 90 41 2d
81 00 90 41 00 00 90 40 32 40 90 40 00 40 90 40 28 40
90 40 00 40 90 3e 2d 40 90 3e 00 40 90 3e 32 40 90 3e
00 40 90 3c 1e 82 00 90 3c 00 00 ff 2f 00
```

Organization/Parsing file

- **Beat class**
  - Notes on beat
  - Beat number

    (8)

Notes off beat
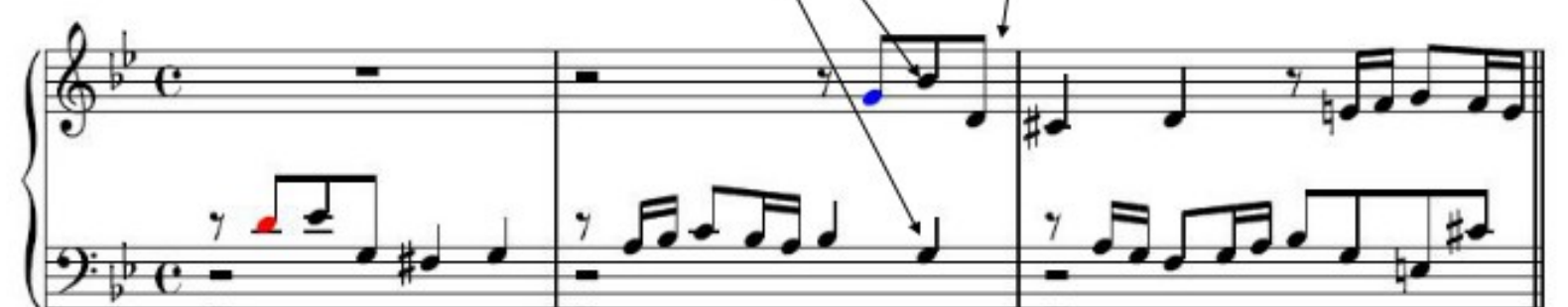
Project Design

Code was written to read and organize the information stored in a midi file. A midi file is a sequence of commands preceded by how long to wait until they are executed (delta-time). For the purposes of this project, I will only need the note-on and note-off commands and their delta-time values. This will allow the code to determine which notes are playing at the same time and what rhythms are found in the piece. This information is stored in a Beat class.

Each instance of a Beat knows which notes sound on its downbeat and off the beat. It also contains which beat it represents from the piece (beat 1, beat 2, etc.), and can be sorted according to this number. When the notes that sound within a beat are passed to a chord identification method, it returns what chord most accurately represents the notes in that beat.

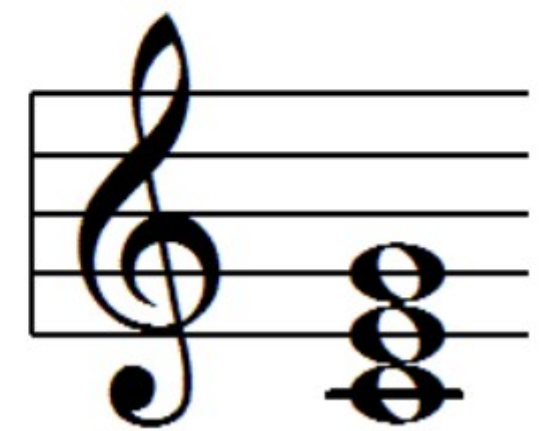## Chord Identification

Notes: C, E, G

What kind of chord? Look at intervals

E: m3, m6

G: P4, M6

C: M3, P5

- These form

a major chord

Discussion

This project could be a valuable tool in organizing libraries of music. Genres are often generalities that do not necessarily fit a piece exactly, but this project could determine what music really is similar and should be grouped together. It could also provide interesting insights into how we interpret music and what gives different styles their distinct feel.

Another application could be composer identification. Sometimes, new manuscripts of music are discovered, but their composer is not known. Successful implementation of this project could aid in identifying the composer of any newly discovered composition.