Musical Genre Similarity

Presented by:

Stephen Sinclair

MUMT 611 Winter 2006, McGill University

Uses

- Recommendation engines - commercial potential.
- Personal music databases organization.
- Large-scale music database searching by genre.

Problems

- Genre is not well defined. (Aucouturier and Pachet 2003)
- Social context:
  - Where was a song produced?
  - Who are the artists friends with?
  - In what era or decade was it produced?
- However, we can assume genre identifies music that "sounds" similar.
  - Instrumentation
  - Melodic structure
  - Rhythmic structure

Matching a taxonomy

- Superimposing onto a taxonomy
  - “Prescriptive” approach
- Emerging a taxonomy from the data

Features

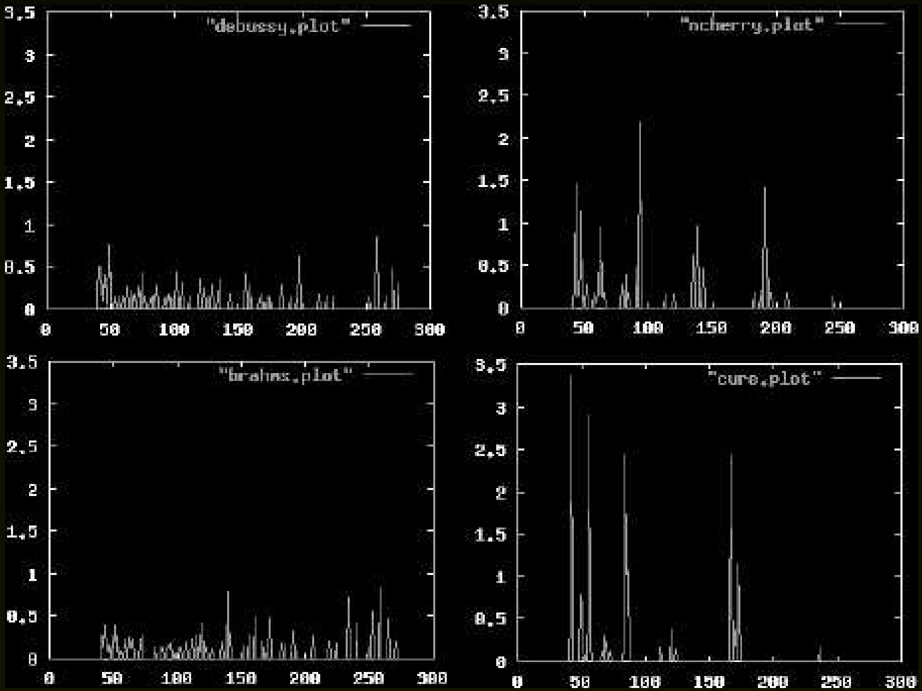
- Timbre-related
- Rhythm-related
- Pitch-related

Timbre-related Features

- FFT coefficients
- Cepstrum
- Mel cepstrum coefficients (MFCC)
  - Non-linear perceptual frequency scale
- Linear prediction
- MPEG filterbank components
- Spectral centroid
- Spectral flux
- Zero-crossing rate
- Spectral roll-off
- Low-order statistics
- Delta coefficients

Rhythm-related Features

- Beat histogram (Tzanetakis et al. 2001)
- Good for differentiating between:
  - Simple rhythms (Rock, Pop)
  - Complex rhythms (World music)
  - Subtle percussion (Classical)



Pitch-related Features

- Less-often used
- Pitch histogram
  - Rock tends to have histogram peaks
  - Jazz histograms are flatter - more notes are played
- Must deal with pitch recognition
  - Typical problems
  - “Songs” are polyphonic signals
- May be useful to restrict range of data to symbolic recordings (e.g., McKay and Fujinaga 2004)

Classification

- User-specified taxonomies are
  - Ambiguous or inconsistant
  - Too small (e.g., rock, jazz, classical.)

Typical results

- It is easy to separate Classical and Techno
- It is difficult to separate Rock, Pop, and Country.
    - Consider instrumentation, timbre.

Perceptual study

- Soltau 1998
  - Compared 37 subjects' ability to classify pop and rock.
  - Human confusions were similar to system's confusions!

Genre-dependant features

- Different genres emphasize different frequency ranges
- Over-specifying features can be detrimental to classification
- Classifying within a genre should ignore features that only serve to separate other genres.
- No guarantee that training set is sufficient to define the best feature set.

Automatic taxonomy

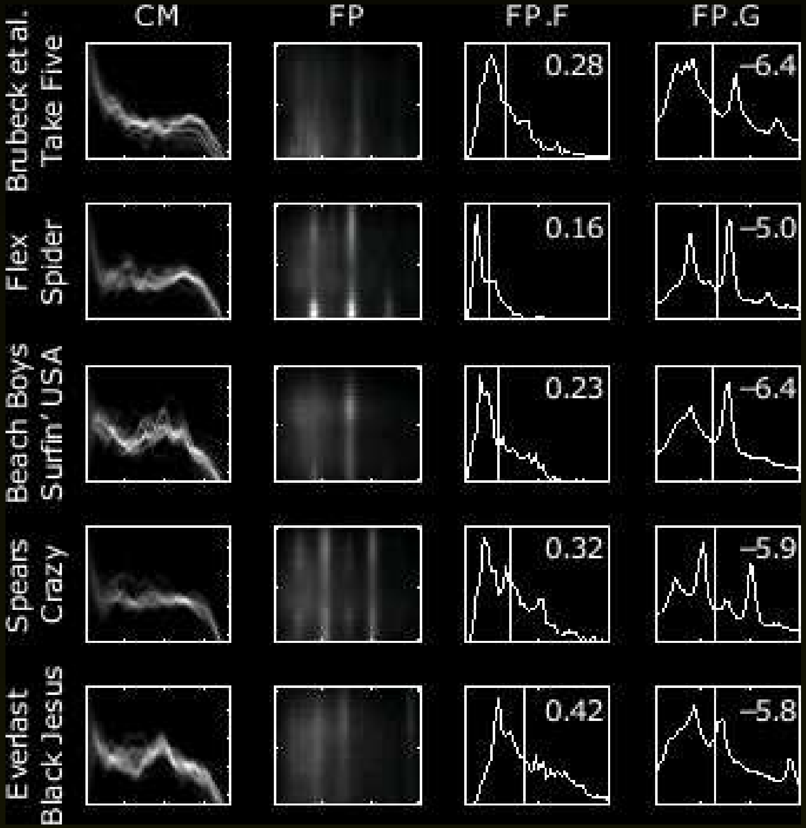
- Clustering a new taxonomy based on similarity measures.
- Problem:
  - Clusters are not labeled

Data-mining techniques

- Collaborative filtering
  - Hits tend to dominate.
  - No guarantee that buying habits are a good indication of genre.
- Data mining (Co-occurrence analysis)
  - Tracklists
  - Compilations
  - Radio program archives

Improvements (Pampalk et al. 2005)

- MIREX '05 submission
- Fluctuation patterns
  - Focus
  - Gravity



Improvements (Pampalk et al. 2005)

- The “Artist filter”
  - An artist should not be in both training and testing sets
  - Drastic lowering effect on the performance of classifier
  - Meaning it *should* be used!
  - Otherwise we are classifying artist, not genre.

References

Aucouturier, J.-J. and F. Pachet (2003). Representing musical genre: A state of the art. *Journal of New Music Research 32*(1).

McKay, C. and I. Fujinaga (2004). Automatic genre classification using large high-level musical feature sets. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR)*.

Pampalk, E., A. Flexer, and G. Widmer (2005, September). Improvements of audio-based music similarity and genre classification. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR)*.

Soltau, H. (1998). Recognition of musical types. In *Proceedings International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.

Tzanetakis, G., G. Essl, and P. Cook (2001). Automatic musical genre classification of audio signals. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*.