PHILIPS

Music Content Analysis

Martin F. McKinney
Research/DSP
Audio Content Analysis
February 20, 2007

Overview

- Introduction
- Research at NatLab
  - Audio Classification
    - General Audio
    - Music Genre
    - Music Mood
    - Discrimination of low-quality audio/music
  - Feature Development
  - Applications
- Demo: Audio Classification

Why analyze audio content?

- Consumers have access to a growing amount of multimedia
  - Storage capacity is increasing
  - Transmission bandwidths are increasing
  - Compression techniques are improving

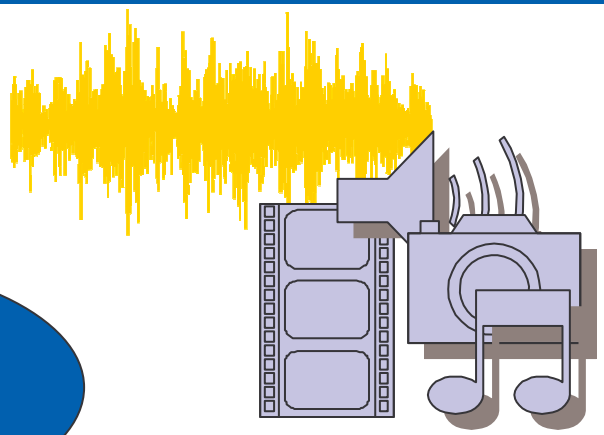
Consumers need descriptors of audio

- To organize, explore and navigate their music databases and other sources.
- Content evaluation vs. annotated database
  - Automatic Content Evaluation
    - Pros
      - Evaluation can be done offline or “on the go”
      - Evaluation scale is continuous (not discrete)
      - Analysis can be locally specific
    - Cons
      - Certain information is difficult to extract from audio
  - Annotated database (e.g., gracenote, moodlogic, allmusic)
    - Pros
      - Variety of metadata is unlimited
    - Cons
      - Not always available
      - Not always reliable or consistent with other databases

State of the art

- Recent explosion of interest in automatic audio and music analysis
  - Increased activity at conferences: ISMIR, DAFX, ICME, ICASSP
- Typical approach: feature extraction followed by pattern classification
- MPEG7
  - Description schemes for multimedia material
  - Reference software for automatic analysis
  - http://www.chiariglione.org/mpeg/

General approach



Feature Extraction

Classification

Segmentation

Comparison

Music, speech, noise, crowd

Music genre

Music mood

Mark speaker changes

Mark radio music vs. non-music

Chapter programs and commercials

Play similar or different music

Find same speaker

User preference profiling

Literature Search: Findings

- Typical method: Feature extraction followed by pattern classification
- Specific method of classification is not always crucial
  - i.e., features are the limiting factor
- Temporal properties of audio are important for classification and summarization
- Our concluding goals:
  - Focus on improving features
  - Describe temporal properties of features

Feature Comparison: Experiment

- Compare classification performance of four feature sets:
  - "Standard" low-level signal parameters
  - Mel-frequency cepstral coefficients (MFCC)
  - Features based on models of auditory perception
- Include statistics of feature temporal behavior as additional features

Feature Comparison: Conclusions

- Classification based on features from perceptual models is better than that from other standard feature sets.
- Temporal modulations of features are important for audio and music classification.
- Feature development can improve audio and music classification.

Audio Feature Extractor for Music

- Developed for music applications: genre and mood classification, LikeMusic
- AFX Components:
  1. Spectrotemporal features
     - McKinney, M.F., Breebaart, D.J., Features for audio and music classification, ISMIR, Oct. 2003.
  2. Percussiveness features
     - Skowronek, J. and McKinney, M.F., New features for music classification: Percussiveness of sounds, in *Intelligent Algorithms in Ambient and Biomedical Computing*, Verhaegh, W.F.J., Aarts, E.H.L. and Korst, J. (eds), 2006.
  3. Tonality features
     - Par, S.L.J., McKinney, M.F., Redert, P.A., Musical key extraction from audio using profile training, ISMIR, Oct 2006.
  4. Rhythm features
     - McKinney, M.F., and Moelants, D., Ambiguity in tempo perception: What draws listeners to listeners to different metrical levels?, *Music Perception*, 24(2), 2006, pp. 57-70.

Applications

- Music Mood Classification
  - Skowronek, J., McKinney, M.F., Par, S.L.J., Ground truth for automatic music mood classification, ISMIR, Oct 2006.
- Music Similarity Assessment
  - Novello, A., McKinney, M.F., Kohlrausch, A.G., Perceptual evaluation of music similarity, ISMIR, Oct 2006.
- Music Structure Analysis
  - Bruderer, M. Kohlrausch, A.G., and McKinney, M.F. Perception of structural boundaries in popular music, International Conference on Music Perception and Cognition, Bologna, 2006.

Demo

- Audio Classification


PHILIPS