Combining Audio and Symbolic Descriptors for Audio Music Classification and Similarity Retrieval

Thomas Lidy, Andreas Rauber  
Vienna University of Technology, Austria  
Department of Software Technology and Interactive Systems  
www.ifs.tuwien.ac.at/mir

Antonio Pertusa, José Manuel Íñesta  
University of Alicante, Spain  
Departamento de Lenguajes y Sistemas Informáticos  
http://grfia.dlsi.ua.es/cm

System Overview

Audio Features
Rhythm Histogram (RH)  
rhythmic energy for 60 modulation frequency bins (60 dim.)

Statistical Spectrum Descriptors (SSD)  
statistical measures describing fluctuations on Bark-scale Sonogram (167 dim.)

Onset Features  
statistical measures of onset strength & intervals (11 dim.)

Transcription
Polyphonic Transcription System: Simple and fast multi-f0 estimator
Only frames after onsets are analyzed to get notes

Algorithm:
Sinusoidal Likeness Measure (SLM)
- Spectral peaks are F0 candidates if:
  - amplitude is higher than a given threshold
  - frequency is within the range [50-1200] Hz
  - close to the Western pitches
Harmonics are identified
Constant spectral pattern is subtracted
Removal of F0 candidates according to thresholds

Symbolic Features
extracted from transcribed MIDIs
statistics of:
- note pitches
- note durations
- number of notes
- number of non-significant silences
- number of distinct pitch intervals
- Inter Onset Intervals etc.

feature selection based on previous studies on symbolic genre classification

37 features in total

MIREX 2007 Results
Genre Classification  Mood Classification  Artist Identification  Composer Id.  Audio Music Similarity and Retrieval

presented at
ISMIR 2007  
8th International Conference on Music Information Retrieval  
September 23 – 27, 2007 • Vienna, Austria