Transcription and feature-based analysis of the electric bass guitar - Application for music and artist classification

Jakob Abeßer (jakob.abesser@idmt.fraunhofer.de)
Fraunhofer IDMT, Germany

Fraunhofer IDMT

Introduction

Motivation
- Analyzing music requires a profound view on both
  - the musical piece itself (e.g. composition, genre)
  - the performing musicians (e.g. style, timing, articulation)
- Describing the semantics of music performances is suited for high-level music characterization due to the close relation to musicology
- Instrumental solo parts offer the biggest freedom for an individual musical expression to a musician

Goals
- Extension of common bass transcription methods to reveal and model
  - Playing styles (genre-specific, e.g. Walking Bass)
  - Plucking & expression styles (instrument-specific, e.g. Slap)
- Development of transcription-based high-level features & models to describe different aspects of musical improvisation
  - Focus on different musical domains (e.g. rhythm, melody)
- Evaluation within genre and artist classification scenarios
  - Focus on instrumental solo parts (improvisation)

Challenges
- Extraction of score parameters from polyphonic and multi-timbral audio data
- Interdependence between the meaningfulness of the calculated high-level and precise and complete transcription results
- Artists playing the same instruments within related musical genres are difficult to distinguish even to experienced listeners
- Modeling musicological knowledge about instrument– and genre-specific playing styles (e.g. Walking Bass) to make them retrievable

Feature-based analysis [3, 4]

Pre-processing
- Extraction of score parameters (note pitch, onset, duration)
  - Symbolic audio data (MIDI) : **MIDI toolbox for MATLAB** [1]
  - Real audio data (MP3) : **Transcription Toolbox** [2]
    - Software toolbox encapsulating four different transcription algorithms for 4 instrument groups (melody, harmony, bass, and drum instrument)
    - Automatic extraction of the beat grid enables a projection of all note onsets from their values in milliseconds to multiples of bar lengths

Feature Extraction
- **Melody & Harmony**
  - Derived from the absolute pitch (pitch range, ratio of constant note sequences, chord tone ratio)
  - Derived from the relative pitch (measure of chromatics, measure of sequences with a constant interval direction, dominant direction)
  - Occurance of different interval types (primes, seconds etc.)
- **Rhythm**
  - Measure of syncopation & swing factor (different grids)
  - Dominant rhythmic grid (4/ 8/ 16…), feeling (down-/ off-beat), characteristic (binary, ternary)
  - Rhythmic precision (as an inverse measure of "quantization costs" towards a certain grid)
- **Structure / Repetitions**
  - Retrieval of rhythmic and melodic repetitions within instrumental tracks
  - Application of a pattern search algorithm for character strings derived from
    - Absolute pitch
    - Quantized onset & duration
  - Features derived from statistical properties of length, incidence rate and mean distance of the detected patterns
- **Interaction**
  - Features characterizing the rhythmical similarity between the tracks
  - Chord-tone ratio in the melody as a measure of harmony-relatedness
  - Measure of „doubled" notes between bass & bass-drum
- **Further bass-related features**
  - Dominant bass pattern, measure of tonal & rhythmic variation of the bass

Evaluation
- Excerpts from instrumental solo parts (25s - 40s)

- **Genre classification** [3, 4]
  - MIDI : 50 tracks / genre, Audio : 40 excerpts / genre
  - 6 genres (Swing, Blues, Funk, Latin, Metal-Hardrock, Pop)
  - **Experiment 1** [3]
    - 4x148 = 592 high-level features, 4 instruments (Melody, Harmony, Bass, Drums)
    - Best results (84% MIDI, 63.4% Audio), LDA(5), SVM(RBF)
  - **Experiment 2** [4]
    - 154 bass-related high-level features
    - Best results (81% MIDI, 46% Audio), IRMFSP(40), GDA

- **Artist classification** [3]
  - 2 sets à 4 artists, 30 excerpts / artist
    - E-Guitar (E. Clapton, R. Gallagher, J. Hendrix, S. R. Vaughan)
    - Saxophone (J. Coltrane, D. Gordon, C. Parker, J. Redman)
  - Results: E-Guitar (58.8%), Saxophone (56.0 %)

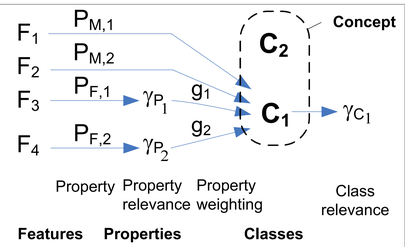
- **Future Work**
  - Evaluation based on larger data-sets
  - Further genres / instruments

Concept-Class-Model [4]

Goal
- Generic Framework to translate known musicological properties into explicit restrictions on feature values
- Implementation of a concept-based classifier for genre classification

Advantages
- Significantly fewer features are necessary to model each class as is in common machine learning approaches
- Approach is based on rules (expert system) and is closely related to musicology

Overview

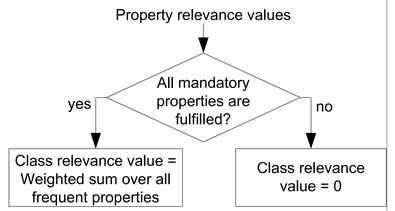


Concept
- The term generally represents all approaches to categorize music
- Examples: BassPlayingStyle (SB), Genre (GE)

Class
- Instances of a concept
- Defined / characterized by different properties
- Examples: WalkingBass (concept SB), Swing (concept GE)
- The *class relevance value* γ_C quantifies to what extend a class is relevant for the musicological description of a song ([0,1])

Property
- Correspond to restrictions on feature values (such as "isBiggerThan" or "isEqual")
- Different types
  - **mandatory** (strictly need to be fulfilled) <-> **frequent** (not compulsory, assigned with a weighting factor g_i)
  - **omnipresent** (constantly valid) <-> **conditional** (depend on a certain condition)
- A *property relevance value* γ_P quantifies to what extend a property is fulfilled (γ_P = 1) or not (γ_P=0)

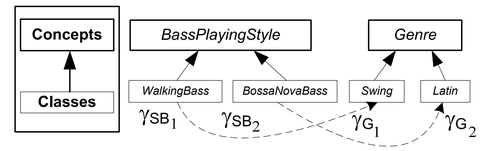
Algorithm



Example
(Excerpt)

Class „WalkingBass" | **Concept** „BassPlayingStyle"
(1) A frequent use of chord tones is mandatory.
$P_{1,MO}$: F { *ChordToneRatio* } isBiggerThan 0.3
(2) The melodic direction is often constant within each bar. (important property - weighting factor g_2 = 0.7)
$P_{2,FO}$: F { *ConstantDirection* } isBiggerThan 0.7
(3) If quarter notes are primarily used (such as in slow and mid-tempo Jazz songs), there is a high swing factor related to the eighth note grid. (important property - weighting factor g_3 = 0.8)
If Condition: F { *DominantGrid* } isEqual 4 $P_{3,FC}$: F { *SwingFactor , 8* } isBiggerThan 0.7
(4) If eigth notes are primarily used (such as in up-tempo Jazz songs), there is a high swing factor related to the sixteenth note grid. (important property - weighting factor g_4 = 0.8)
If Condition: F { *DominantGrid* } isEqual 4 $P_{4,FC}$: F { *SwingFactor , 16* } isBiggerThan 0.7

Evaluation
- Genre classification based on detected bass playing styles



- Symbolic audio data, 50 excerps (20-40s) / genre
- 6 classes of the concept „BassPlayingStyle" defined
  - **WalkingBass** (Swing) - 42%
  - **BluesShuffle** (Blues) - 68%
  - **FunkSyncopated** (Funk) - 46%
  - **SteadyRiff** (MetalHardrock) - 34%
  - **BossaNovaBass** (Latin) - 70%
  - **ChordRootAccompaniment** (Pop) - 6%
- 5 Properties per class
- Property weightings & thresholds were derived from an development data set
- Mean accuracy: 44.3%, results strongly vary throughout the testset (6.0% - 70.0%), proof of concept provided

Future steps
- Defininning more properties per class
- Definition of further styles for each genre
- Weighted allocation of styles towards multiple genres

Transcription

Goal
- Extraction of bass-related
  - **Plucking Styles** (Finger style, Muted, Pick, SlapThumb, SlapPluck)
  - **Expression Styles** (Normal, Vibrato, Bending, Harmonics, Dead notes)

Method
- Various low– and mid-level audio features
- Evaluation of both the attack and decay phase of single notes
- Temporal integration methods
- Feature selection & feature space transformation (IRMFSP, LDA)
- Compare different classifiers (SVM, GMM, HMM)

Future steps
- Combine „common" transcription methods (pitch, velocity, onset, duration) with the detection of both plucking & expression styles (based on low– and mid-level features) and playing styles (based on high-level features)

Selected Publications

[1] Eerola, T., Toiviainen, P. *„MIDI toolbox: Matlab tools for music research"*. In www.jyu.fi/musica/miditoolbox (last call: 09.10.2008), Jyväskylä, Finland, 2004, University of Jyväskylä
[2] Dittmar, C., Dressler, K., Rosenbauer, K. *„A toolbox for automatic transcription of polyphonic music"*. In Proc. of the Audio Mostly, 2007
[3] Jakob Abeßer, Christian Dittmar, Holger Grossmann. *Automatic genre and artist classification by analyzing improvised solo parts from musical recordings.* In: Proceedings of the Audio Mostly, Piteå, Sweden, 2008
[4] Jakob Abeßer, Hanna Lukashevich, Christian Dittmar, and Gerald Schuller. *Genre classification using bass-related high-level features and playing styles*. In: Proceedings of the ISMIR, 2009