Diplomarbeitspräsentationen der Fakultät für Informatik

Discriminant Analysis of Three Rhythmic Descriptors in Musical Genre Classification

Bernhard Pflugfelder
Bernhard.Pflugfelder@gmail.com

Masterstudium:
Computational Intelligence

Information & Software Engineering Group

Technische Universität Wien
Institut für Softwaretechnik und interaktive Systeme
Arbeitsbereich: Information and Software Engineering
BetreuerIn: Ao. Univ.-Prof. Dr. Andreas Rauber
sowie DI Thomas Lidy

Motivation

... provide an intuitive understanding for categorization of musical pieces

Musical Genres ...

Hip hop, Classical, Jazz, Rock, Country, Blues, R&B, World, Electronic

... are frequently used by humans to categorize music collections, e.g. music retailers or music libraries

... are widely used in *Music Information Retrieval* (MIR) to organize digital music archives for effective access and search of musical pieces

It is generally assumed that genres posses an intrinsic descriptive power to constitute specific *musical characteristics* [1]

This means that a genre may be uniquely related to a specific ...


Rhythm and

Melody and

Instrumentation and others...

Goals & Applications

Two key goals were examined:

1st Goal: Evaluate discriminative rhythmic feature patterns in order to distinguish musical genres.

... based on the rhythmic descriptors Rhythm Patterns, Statistical Spectrum Descriptor & Rhythm Histogram

2nd Goal: Evaluate the usefulness of dimensionality reduction based on the discriminative power of every feature.

Sub goal: Design of the *DiscriminationAnalyzer* application

Applications: rhythmic genre descriptions, feature selection, Hierarchical Genre Classification, ...

Discriminant Analysis

Idea

- Genres provide a discrimination power due to the rhythmic aspect of music

  Question: *"Are genres related to specific rhythmic elements?"*, e. g.:
  - specific range of beats per minute (BPM)
  - key frequencies bands

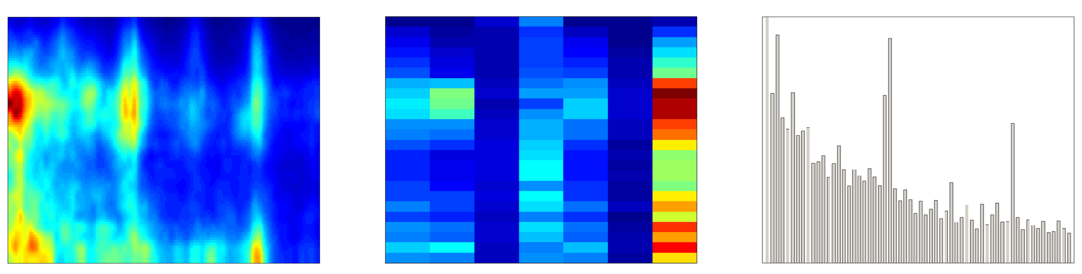
Concept

- Use statistical variable dependency to constitute genre discrimination of a feature
- Consider two different approaches:
  - Mutual Information
  - Value-based Class Determination with Nearest-Neighbor algorithm
- Employ five heuristic models:
  - Chi-square and Information Gain
  - Gain Ratio and Balanced Information Gain
  - ReliefF

Rhythmic descriptors

- Analyze three different rhythmic descriptors [2]:

Rhythm Patterns (RP) | **Statistical Spectrum Descriptor (SSD)** | **Rhythm Histogram (RH)**



- Time-invariant representation
- All descriptors use psycho-acoustic transformation
- 1440 features for RP:
  - 24 critical bands x 60 modulation frequencies
- 168 features for SSD:
  - 7 statistical moments x 24 critical bands
- 60 features for RH:
  - Modulation frequencies are grouped into 60 bins

Computation

- Apply heuristic models for each descriptor & genre
- Establish *one-vs.-rest* labellings
- Robust estimation due to multiple fold computation
- Fold result verification by testing with the Kendall's rank correlation coefficient
- Aggregate final results by averaging

DiscriminationAnalyzer

Idea

- Combine tools for discriminant analysis along with feature selection and feature subset evaluation

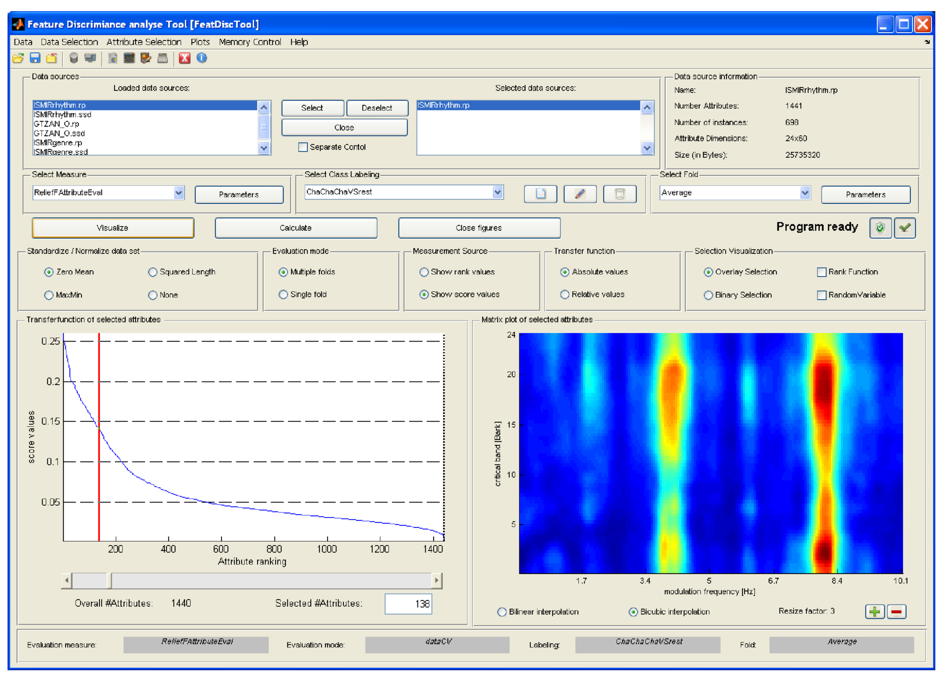


Fig. 1 *Main window of DiscriminationAnalyzer*

Key properties

- Arbitrary feature sets usable
- Simultaneous processing of loaded sets
- Includes 7 selectable heuristic models
- Interface to integrate user-defined models
- Visual and numeric result representation
- Interactive feature selection and subset evaluation

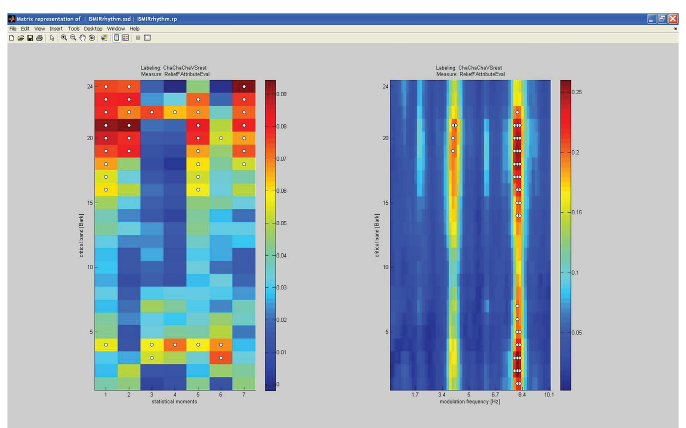


Fig. 2 *Visualization of discriminative features based on SSD (left) and RP (right)*

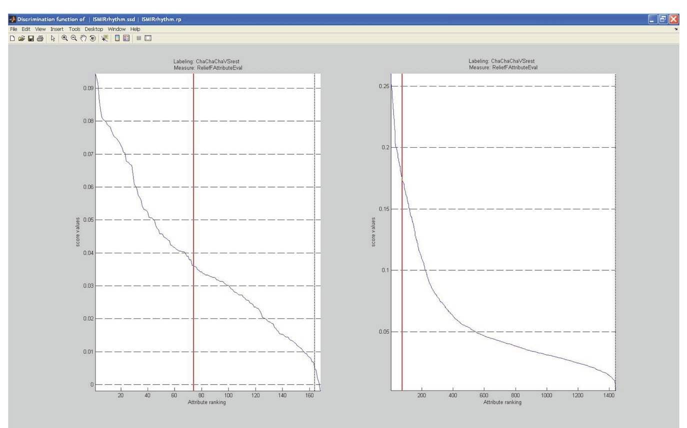


Fig. 3 *Discrimination functions according to results of figure 2. The red bars help to select most discriminative features.*

Input & Output

- ARFF dataset format of WEKA [3]
- SOMLib dataset format (+ ground truth)
- MAT format for computation persistence

Dimensionality Reduction

Observation

- Large feature sets can cause deteriorating classification performance
  - *curse of dimensionality*

Idea

- Select a feature subset according to the discriminative power of every feature
- Question: *"Does the classification performance change due to this dimensionality reduction? And if yes then how?"*

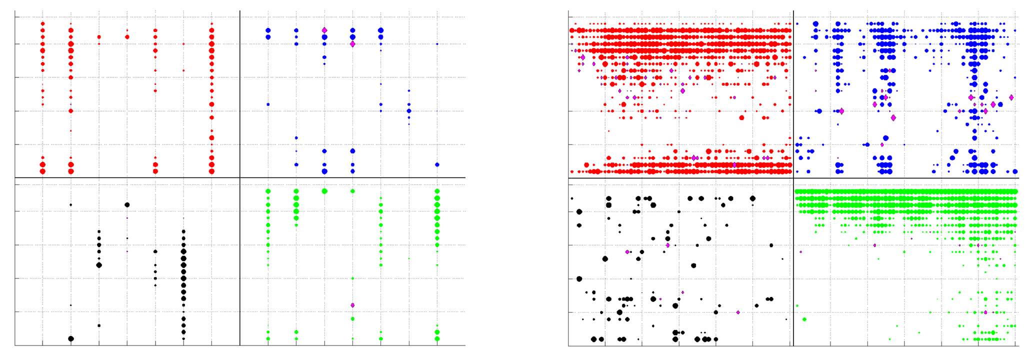


Fig. 4 *Exemplary feature subsets according to SSD (left) and RP (Right) containing 50 % of most discriminative features only.*

Evaluation setup

- Subsets of *k* most discriminative features
- Use 30 linearly distributed samples for *k*
- Evaluation methodology:
  - One-vs.-rest labellings for every genre
  - 10-fold cross validation
- Three learning models:
  - Support Vector Machine (SMO)
  - Decision Tree (J48)
  - Naive Bayes
- Use of WEKA workbench [3]

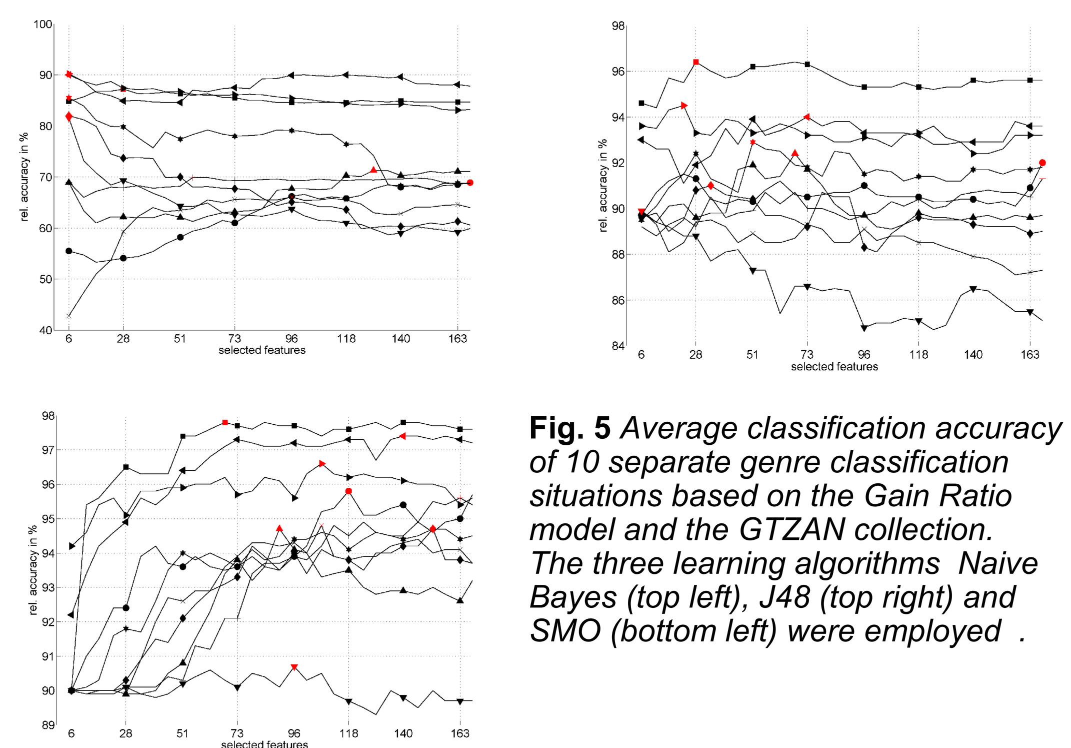


Fig. 5 *Average classification accuracy of 10 separate genre classification situations based on the Gain Ratio model and the GTZAN collection. The three learning algorithms Naive Bayes (top left), J48 (top right) and SMO (bottom left) were employed .*

Conclusions

Benchmark music collections

- Three music collections were used for both evaluations:
  - GTZAN [4]
  - ISMIR 2004 Genre [5]
  - ISMIR 2004 Rhythm [5]

Collection Name	Genres	Samples
GTZAN	10	1.000
ISMIR 2004 Genre	6	1.458
ISMIR 2004 Rhythm	8	698

Discriminant analysis:

- Diverging feature patterns for all genres according to all music collections
- Individual feature patterns according to various genres
- Calculation models based on the impurity function performed quite consistently.
- SSD: variance and skewness irrelevant

A similar performance of the three calculation models could not be concluded for all descriptors. Highest degree of similarity in the case of the SSD.

Effectiveness of the feature selection approach:

- Results slightly varied according to Chi-square, IG, GR, Balanced IG, where GR & Balanced IG should be preferred.
- Accuracy was limited by a margin of ~ 5 % with some stronger variations.
- For J48 and SMO, a margin of 1 – 2 % was concluded for almost all genres.
- A margin of 1 – 2 % was generally concluded when 50 % or more of the most discriminative features were used.

The effectiveness of the feature selection approach could be definitely concluded.

Future Work

- Use within with "real-world" genre classification systems:
  - Hierarchical Classification
  - Ensemble classification
- Discriminative feature weighting and subset selection

References

[1] Jean-Julien Aucouturier and François Pachet. Representing Musical Genre: A State of the Art. Journal of New Music Research, 32(1):83-93, 2003.

[2] Thomas Lidy and Andreas Rauber. Evaluation of Feature Extractors and Psycho-Acoustic Transformations for Music Genre Classication. In Proceedings of the International Conference on Music Information Retrieval (ISMIR 05), pages 34-41, 2005.

[3] Ian H. Witten and Eibe Frank. Data Mining: Practical machine learning tools and techniques. Information Science and Statistics. Morgan Kaufmann, US, 2nd edition, 2005.

[4] George Tzanetakis. Manipulation, analysis and retrieval systems for audio signals. PhD thesis, Princeton University, Princeton, NJ, USA, 2002.

[5] ISMIR 2004 Audio Description Contest. http://ismir2004.ismir.net/ISMIR_Contest.html.