Workshop HfM Karlsruhe

Music Information Retrieval

Classification & Clustering

Christof Weiß, Frank Zalkow, Christian Dittmar, Meinard Müller
International Audio Laboratories Erlangen

christof.weiss@audiolabs-erlangen.de
frank.zalkow@audiolabs-erlangen.de
meinard.mueller@audiolabs-erlangen.de
Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
483 p., 249 illus., hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de
Book: Fundamentals of Music Processing

Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
483 p., 249 illus., hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de
Dissertation: Tonality-Based Style Analysis

Christof Weiß
Computational Methods for Tonality-Based Style Analysis of Classical Music Audio Recordings
Dissertation, Ilmenau University of Technology, 2017

Chapter 7: Clustering and Analysis of Musical Styles
Chapter 8: Subgenre Classification for Western Classical Music
Recall: Chroma Features

L. van Beethoven, 
_Fidelio, Ouvertüre_, 
Slovak Philharmonic
Recall: Chroma Features

- **Orchestra**
  - L. van Beethoven, *Fidelio*, Ouvertüre, Slovak Philharmonic

- **Piano**
  - *Fidelio*, Ouvertüre, arr. Alexander Zemlinsky
  - M. Namekawa, D.R. Davies, Klavier vierhändig
Recall: Chroma Features

- Piano

Fidelio, Overture, arr. Alexander Zemlinsky
M. Namekawa, D.R. Davies, piano four hands
Chroma-based Classification Features

- Chromagram

- Triad features

- Interval features

→ transposition-invariant
Recall: Tonal Complexity

- Realization of complexity measure $\Gamma$
  - Entropy / Flatness measures
  - Distribution over *Circle of Fifths*

\[
\Gamma = \sqrt{1 - r}
\]

- Relating to different time scales!
Recall: Tonal Complexity
Recall: Tonal Complexity

Op. 2, No. 3 Op. 57, No. 1
"Appassionata"

Op. 106, No. 1
"Hammerklavier"

Beethoven Sonatas, 1st movements
Recall: Tonal Complexity
Recall: Tonal Complexity

Haydn, Joseph
1732 – 1809
100 works in dataset
Recall: Tonal Complexity

![Graph showing the evolution of tonal complexity over time with data points for various composers such as Bach, Beethoven, Schubert, and others. The x-axis represents time from 1650 to 2000, and the y-axis represents complexity values ranging from 0.7 to 1.0. The graph includes trends for complexity global, mid-scale, and local.](image-url)
Music Genre Classification

- *world music*
- *JAZZ*
- *HipHop*
- *pop*
- *Rock*
- "classical"

**Baroque**
- J. S. Bach, Brandenburg Concerto No. 2 in F major, I. Allegro, Cologne Chamber Orch.

**Classical**
- L. van Beethoven, *Fidelio*, Overture, Slovak Philharmonic

**Romantic**
- R. Schumann, Sonata No. 2 op. 22, II. Andantino, B. Glemser, Piano

**Modern**
- A. Webern, Variations for Orchestra op. 30, Ulster Orchestra
Music Genre Classification

Subgenre Categories:
- world music
- JAZZ
- HipHop
- pop
- Rock

Period / Era
- Baroque
- Classical
- Romantic
- Modern

Sub-era
- Pre-Classical
- Early Rom.
- Late Rom.

Composer
- Schubert
- Mendelssohn
Music Genre Classification

- Standard approach (*content-based*)
  - Supervised machine learning
  - Based on spectral / timbral features

- In classical music → Instrumentation

- Better categories?
  - *Musical style*
  - Independent from instrumentation
  - → *Tonality / Harmony*
Music Genre Classification

- Typical approach: Supervised machine learning
Music Genre Classification

- Experimental design: Evaluation with Cross Validation (CV)
- Separate data into different parts (*folds*)

- Distribution of classes balanced for all folds
Classification Scenario

- Dataset: CrossEraDB (Historical Periods)
  - Balanced Piano (p) – Orchestra (o)
  - Each 200 pieces → 1600 in total
## Classification Features

<table>
<thead>
<tr>
<th>Standard</th>
<th>Dim.</th>
<th>Tonal</th>
<th>Dim.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFCC</td>
<td>16</td>
<td>Interval cat.</td>
<td>6 x 4</td>
</tr>
<tr>
<td>OSC</td>
<td>14</td>
<td>Triad types</td>
<td>4 x 4</td>
</tr>
<tr>
<td>ZCR</td>
<td>1</td>
<td>Complexity</td>
<td>7 x 4</td>
</tr>
<tr>
<td>ASE</td>
<td>16</td>
<td>Chord progr.</td>
<td>11 x 5</td>
</tr>
<tr>
<td>SFM</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCF</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LogLoud</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NormLoud</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>119</td>
<td>Sum</td>
<td>123</td>
</tr>
<tr>
<td>Mean &amp; Std</td>
<td>x 2</td>
<td>Mean &amp; Std</td>
<td>x 2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>238</strong></td>
<td><strong>Total</strong></td>
<td><strong>246</strong></td>
</tr>
</tbody>
</table>
Dimensionality Reduction

- Reduce feature space to few dimensions (prevent **curse of dimensionality**)
- Maximize separation of classes with **Linear Discriminant Analysis (LDA)**
- Using **standard features** (MFCC, spectral envelope, …)
Dimensionality Reduction

- Reduce feature space to few dimensions
- Maximize separation of classes with **Linear Discriminant Analysis (LDA)**
- Using **tonal features** (interval, triad types, tonal complexity, ... 4 time scales)
Dimensionality Reduction

- Reduce feature space to few dimensions
- Maximize separation of classes with **Linear Discriminant Analysis (LDA)**
- Using **tonal & standard features**
Dimensionality Reduction

- Reduce feature space to few dimensions

- Other methods (supervised):
  - (DNN-based) Autoencoder
  - Feature selection

- Other methods (unsupervised):
  - Principal component analysis (PCA)
  - Nonnegative matrix factorization (NMF)
Classification methods

- $k$ Nearest Neighbours (kNN)
Classification methods

- **k Nearest Neighbours (kNN)**

![Scatter plot of Spectral Flatness vs. Spectral Centroid](image)

**L1-Dist. (Manhattan)**

\[
\|\mathbf{e}\|_1 = \sum_{m=1}^{M}|x_m - y_m|
\]

**L2-Dist. (Euclidean)**

\[
\|\mathbf{e}\|_2 = \sqrt{\sum_{m=1}^{M}|x_m - y_m|^2}
\]

**L∞-Dist. (Maximum)**

\[
\|\mathbf{e}\|_\infty = \max(|x_1 - y_1|, \ldots, |x_M - y_M|)
\]

Slides: Christian Dittmar
Classification methods

- Decision Trees (DT)
Classification methods

- Random Forests (RF)
Classification methods

- Gaussian Mixture Models (GMM)

\[ f(x) = \sum_{g=1}^{G} w_g \cdot p(x | \mu_g, \Sigma_g) \]
Classification methods

- Gaussian Mixture Models (GMM)
Classification methods

- Support Vector Machines (SVM)

\[ f(x) = \text{sgn} \left( \sum_{n=1}^{N} \alpha_n t_n k(x_n, x) + b \right) \]
Classification methods

- Deep Neural Networks (DNN)

\[ f(x) = f_K(f_{K-1}(\cdots f_1(x, w_1)\cdots), w_K) \]
Classification methods

- Deep Neural Networks (DNN)
Classification Results

- Gaussian Mixture Model (GMM) classifier, LDA reduction, 3-fold cross validation

<table>
<thead>
<tr>
<th></th>
<th>Full Dataset</th>
<th>Piano</th>
<th>Orchestra</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard features</strong></td>
<td>87 %</td>
<td>88 %</td>
<td>85 %</td>
</tr>
<tr>
<td><strong>Tonal features</strong></td>
<td>84 %</td>
<td>84 %</td>
<td>86 %</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>92 %</td>
<td>86 %</td>
<td>80 %</td>
</tr>
</tbody>
</table>

Classification Results

- Gaussian Mixture Model (GMM) classifier, LDA reduction, 3-fold cross validation

<table>
<thead>
<tr>
<th></th>
<th>Full Dataset</th>
<th>Piano</th>
<th>Orchestra</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard features</strong></td>
<td>87 %</td>
<td>88 %</td>
<td>85 %</td>
</tr>
<tr>
<td><strong>Tonal features</strong></td>
<td>84 %</td>
<td>84 %</td>
<td>86 %</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>92 %</td>
<td>86 %</td>
<td>80 %</td>
</tr>
</tbody>
</table>

*Overfitting??*

Classification Results

- GMM classifier, LDA reduction, 3-fold cross validation

<table>
<thead>
<tr>
<th></th>
<th>Full Dataset</th>
<th>Piano</th>
<th>Orchestra</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard features</td>
<td>87 %</td>
<td>88 %</td>
<td>85 %</td>
</tr>
<tr>
<td>Tonal features</td>
<td>84 %</td>
<td>84 %</td>
<td>86 %</td>
</tr>
<tr>
<td>Combined</td>
<td>92 %</td>
<td>86 %</td>
<td>80 %</td>
</tr>
</tbody>
</table>

Classification Results

- GMM classifier, LDA reduction, 3-fold cross validation
- **No composer filter**

<table>
<thead>
<tr>
<th></th>
<th>Full Dataset</th>
<th>Piano</th>
<th>Orchestra</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard features</strong></td>
<td>87 %</td>
<td>88 %</td>
<td>85 %</td>
</tr>
<tr>
<td><strong>Tonal features</strong></td>
<td>84 %</td>
<td>84 %</td>
<td>86 %</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>92 %</td>
<td>86 %</td>
<td>80 %</td>
</tr>
</tbody>
</table>

- **Using composer filter**

<table>
<thead>
<tr>
<th></th>
<th>Full Dataset</th>
<th>Piano</th>
<th>Orchestra</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard features</strong></td>
<td>54 %</td>
<td>36 %</td>
<td>70 %</td>
</tr>
<tr>
<td><strong>Tonal features</strong></td>
<td>73 %</td>
<td>70 %</td>
<td>78 %</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>68 %</td>
<td>44 %</td>
<td>68 %</td>
</tr>
</tbody>
</table>

Weiss / Müller, *Tonal Complexity Features for Style Classification of Classical Music*, ICASSP 2015
Classification Results

- GMM classifier, LDA reduction, 3-fold cross validation
- **No composer filter**

<table>
<thead>
<tr>
<th></th>
<th>Full Dataset</th>
<th>Piano</th>
<th>Orchestra</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard features</strong></td>
<td>87 %</td>
<td><strong>88 %</strong></td>
<td>85 %</td>
</tr>
<tr>
<td><strong>Tonal features</strong></td>
<td>84 %</td>
<td>84 %</td>
<td>86 %</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>92 %</td>
<td>86 %</td>
<td>80 %</td>
</tr>
</tbody>
</table>

- **Using composer filter**

<table>
<thead>
<tr>
<th></th>
<th>Full Dataset</th>
<th>Piano</th>
<th>Orchestra</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard features</strong></td>
<td>54 %</td>
<td><strong>36 %</strong></td>
<td>70 %</td>
</tr>
<tr>
<td><strong>Tonal features</strong></td>
<td>73 %</td>
<td>70 %</td>
<td>78 %</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>68 %</td>
<td>44 %</td>
<td>68 %</td>
</tr>
</tbody>
</table>

Weiss / Müller, *Tonal Complexity Features for Style Classification of Classical Music*, ICASSP 2015
Classification Results

- What is actually learned?
- Pay attention to:
  - Overfitting
  - „Curse of dimensionality“ – use dimensionality reduction techniques
  - Artist / album effects
- Evaluation: „Figures of merit“:
  - Confusion matrix
  - Error examples: Consistently misclassified items
  - Listening tests
- Evaluation on unseen data (no cross validation)

Bob Sturm, Classification Accuracy is not enough, Journal of Intelligent Information Systems, 2013
Classification Results – Confusion Matrix

- 80 tonal features, GMM with 1 Gaussian, LDA, composer filtering
- **Full** dataset
- Mean accuracy: 75%
- Inter-class standard deviation: 6.7%
Classification Results: Error Examples

- 80 tonal features, GMM with 1 Gaussian, LDA
- Look at **consistently** and **persistently** misclassified items

<table>
<thead>
<tr>
<th>Class</th>
<th>Composer</th>
<th>Piece</th>
<th>Classified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baroque</td>
<td>Bach, J. S.</td>
<td>Well-Tempered Piano 1, Prelude in Eb minor BWV 853</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Bach, J. S.</td>
<td>Well-Tempered Piano 1, Prelude in F major BWV 856</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Bach, J. S.</td>
<td>Well-Tempered Piano 1, Prelude in A minor BWV 865</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Bach, J. S.</td>
<td>Well-Tempered Piano 1, Prelude in Bb major BWV 866</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Bach, J. S.</td>
<td>Well-Tempered Piano 1, Prelude in Bb minor BWV 867</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Bach, J. S.</td>
<td>English Suite No. 3 in G minor BWV 808, Sarabande</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Bach, J. S.</td>
<td>Brandenburg Conc. No. 1 in F major BWV 1046, Adagio</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Bach, J. S.</td>
<td>Overture No. 2 in B minor BWV 1067, Badinerie</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Bach, J. S.</td>
<td>Overture No. 3 in D major BWV 1068, Gigue</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Couperin, F.</td>
<td>27 Ordres, Huitième ordre, IX. Rondeau passacaille</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Corelli, A.</td>
<td>Concerto grosso op. 6 No. 2, III. Grave – Andante largo</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Lully, J.-B.</td>
<td>Ballet de Xerxes LWV 12, Gavotte en rondeau</td>
<td>Romantic</td>
</tr>
<tr>
<td>Baroque</td>
<td>Purcell, H.</td>
<td>Opera “Dido and Aeneas” Z. 626, Overture</td>
<td>Romantic</td>
</tr>
<tr>
<td>Romantic</td>
<td>Schumann, R.</td>
<td>Kinderszenen op. 15, “Haschemann”</td>
<td>Baroque</td>
</tr>
<tr>
<td>Romantic</td>
<td>Grieg, E.</td>
<td>Holberg suite op. 40, Gavotte</td>
<td>Baroque</td>
</tr>
<tr>
<td>Romantic</td>
<td>Mendelssohn, F.</td>
<td>Symphony No. 4 in A major, IV. Saltarello, presto</td>
<td>Baroque</td>
</tr>
<tr>
<td>Modern</td>
<td>Shostakovich, D.</td>
<td>Preludes &amp; Fugues op. 87 Fugue No. 1 in C major</td>
<td>Baroque</td>
</tr>
<tr>
<td>Modern</td>
<td>Shostakovich, D.</td>
<td>Preludes &amp; Fugues op. 87 Fugue No. 5 in D major</td>
<td>Baroque</td>
</tr>
</tbody>
</table>
Musical Style Analysis
Musical Style Analysis – Complexity
Clustering: Years

- Features: Interval, complexity, chord progressions
- Dimensional reduction with Principal Component Analysis (PCA)
- $k$-means clustering with different number of clusters $k$
Clustering: Pieces

- $k$-means clustering with $k = 5$ clusters

Weiss / Mauch / Dixon / Müller, *Investigating Style Evolution of Western Classical Music: A Computational Approach* 
Musicae Scientiae 2018
Clustering: Composers
Clustering: Composers