INTERNATIONAL AUDIO LABORATORIES ERLANGEN

AUDIO LABS

Workshop HfM Karlsruhe

Music Information Retrieval

Classification & Clustering

Christof Weiß, Frank Zalkow, Christian Dittmar, Meinard Müller

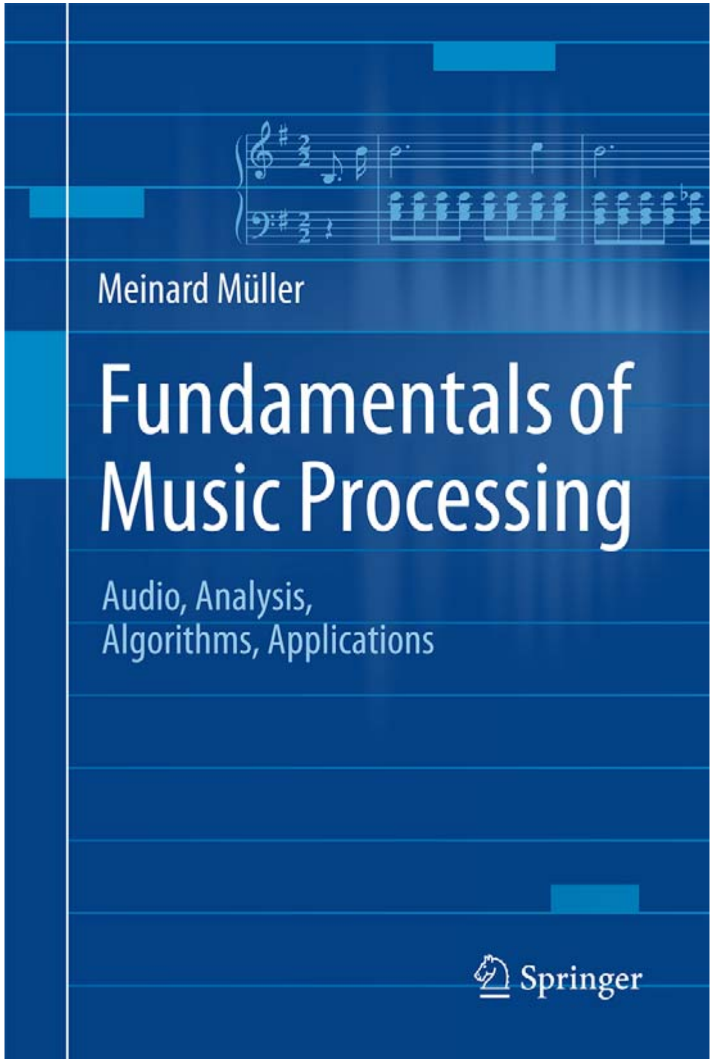
International Audio Laboratories Erlangen

christof.weiss@audiolabs-erlangen.de
frank.zalkow@audiolabs-erlangen.de
christian.dittmar@audiolabs-erlangen.de
meinard.mueller@audiolabs-erlangen.de

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Fraunhofer IIS

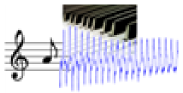
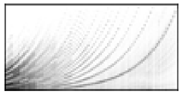
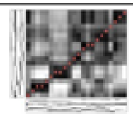
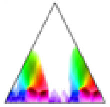
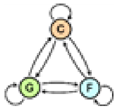
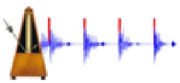
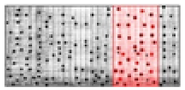
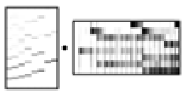
Book: Fundamentals of Music Processing



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

Book: Fundamentals of Music Processing

Chapter		Music Processing Scenario
1		Music Representations
2		Fourier Analysis of Signals
3		Music Synchronization
4		Music Structure Analysis
5		Chord Recognition
6		Tempo and Beat Tracking
7		Content-Based Audio Retrieval
8		Musically Informed Audio Decomposition

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

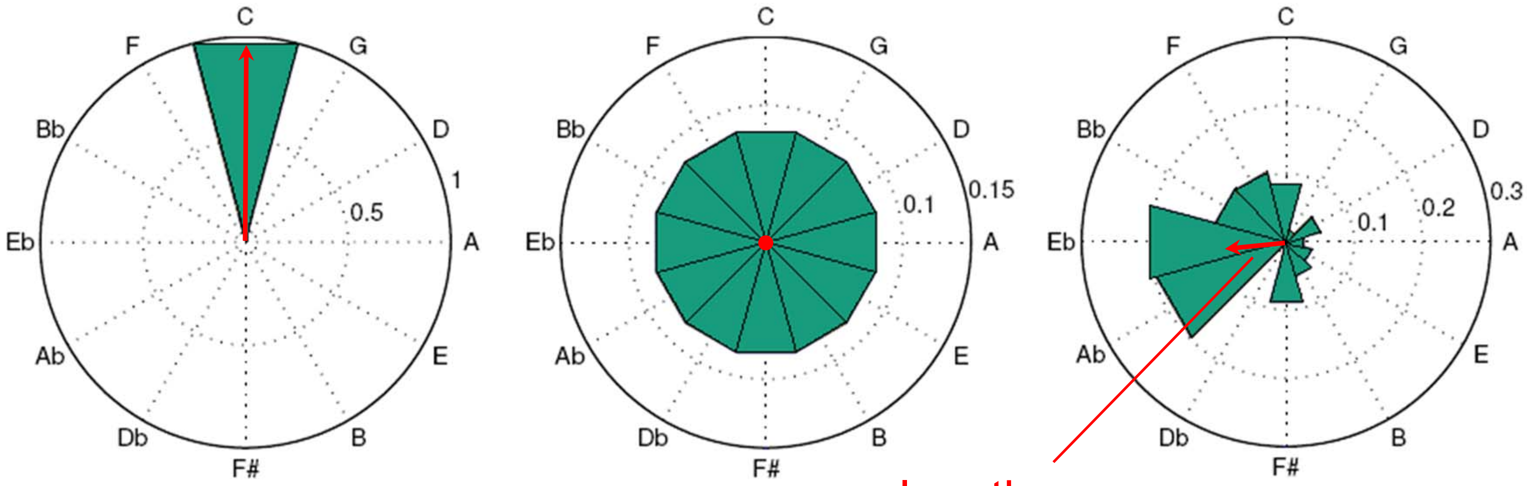
Christof Weiß
Computational Methods for Tonality-Based Style Analysis of Classical Music Audio Recordings
Dissertation, Technical University of Ilmenau 2017
to appear

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

Recall: Tonal Complexity

- Realization of complexity measure Γ
  - Entropy / Flatness measures
  - Distribution over *Circle of Fifths*

$\Gamma = 0$ $\Gamma = 1$ $0 < \Gamma < 1$

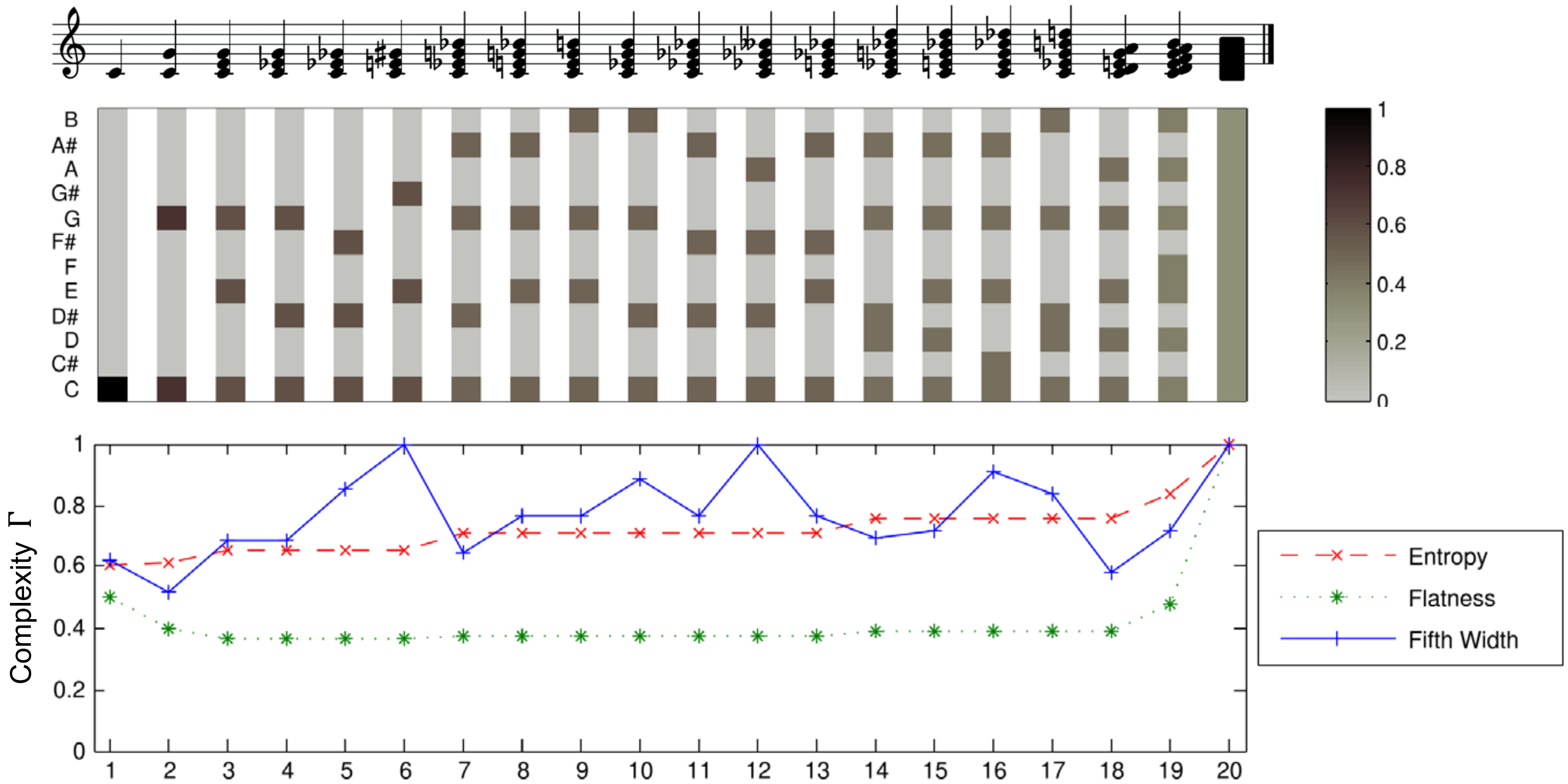


length r

$$\boxed{\Gamma = \sqrt{1 - r}}$$

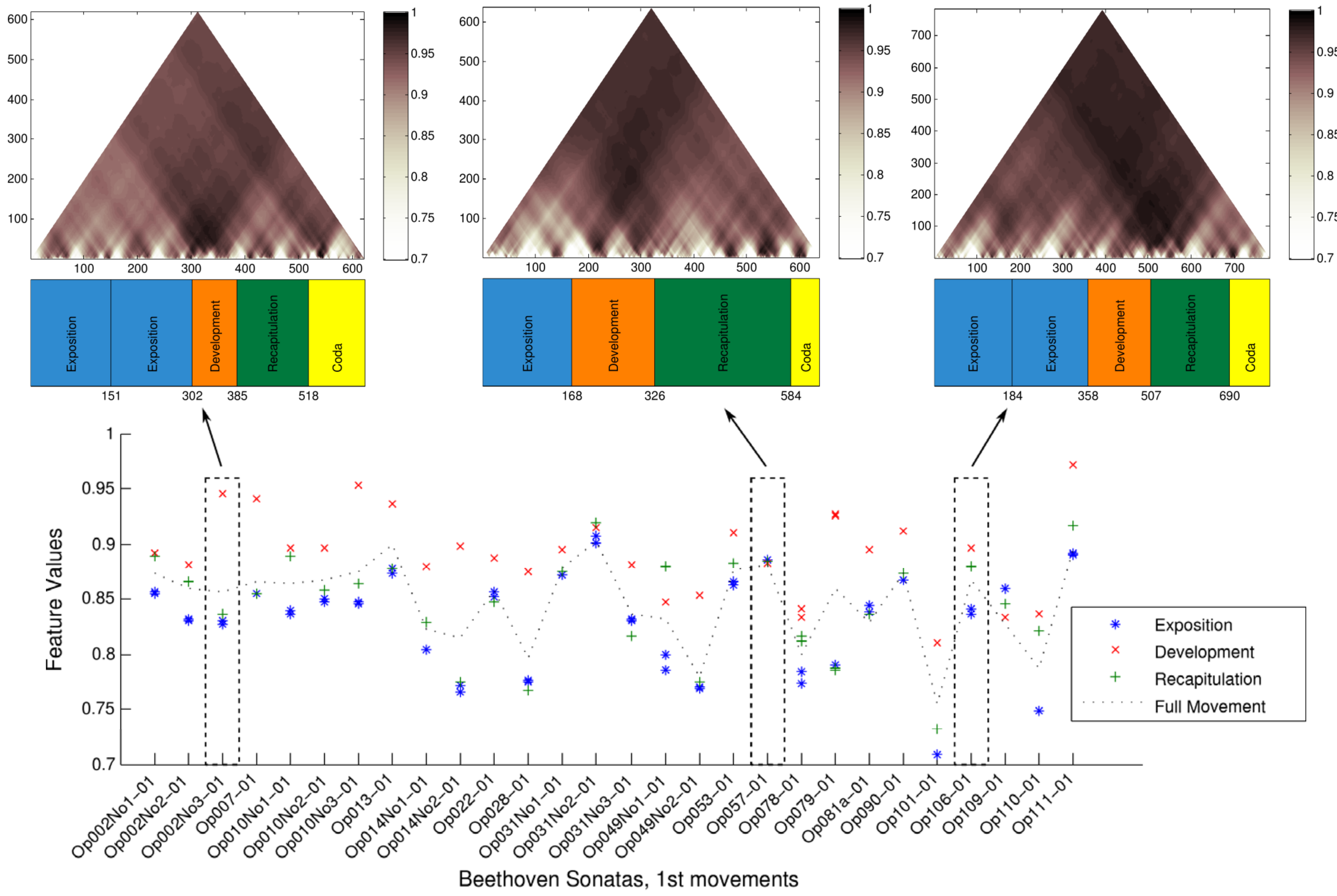
- Relating to different time scales!

Recall: Tonal Complexity

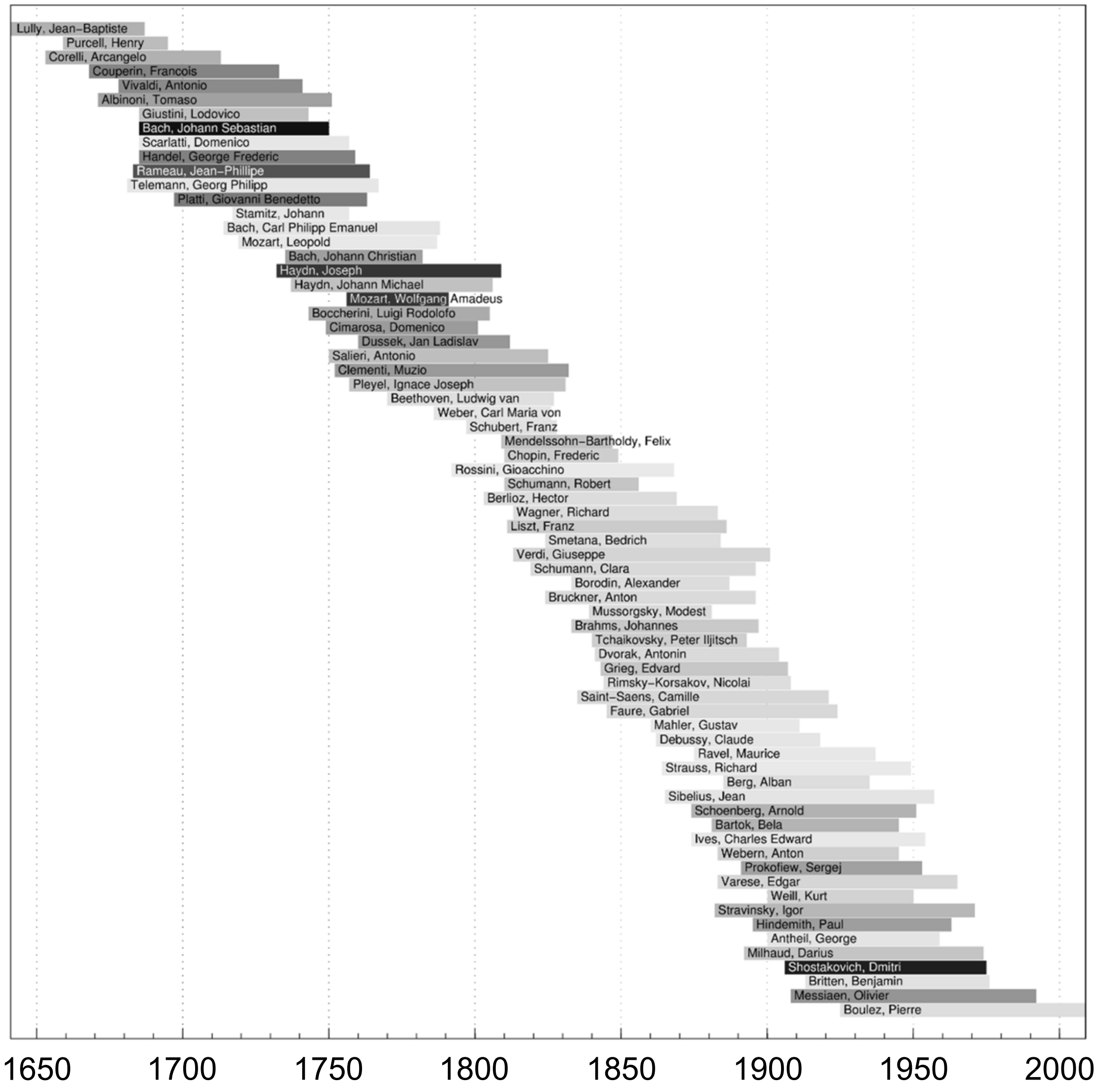


Weiss / Müller, *Quantifying and Visualizing Tonal Complexity*, CIM 2014

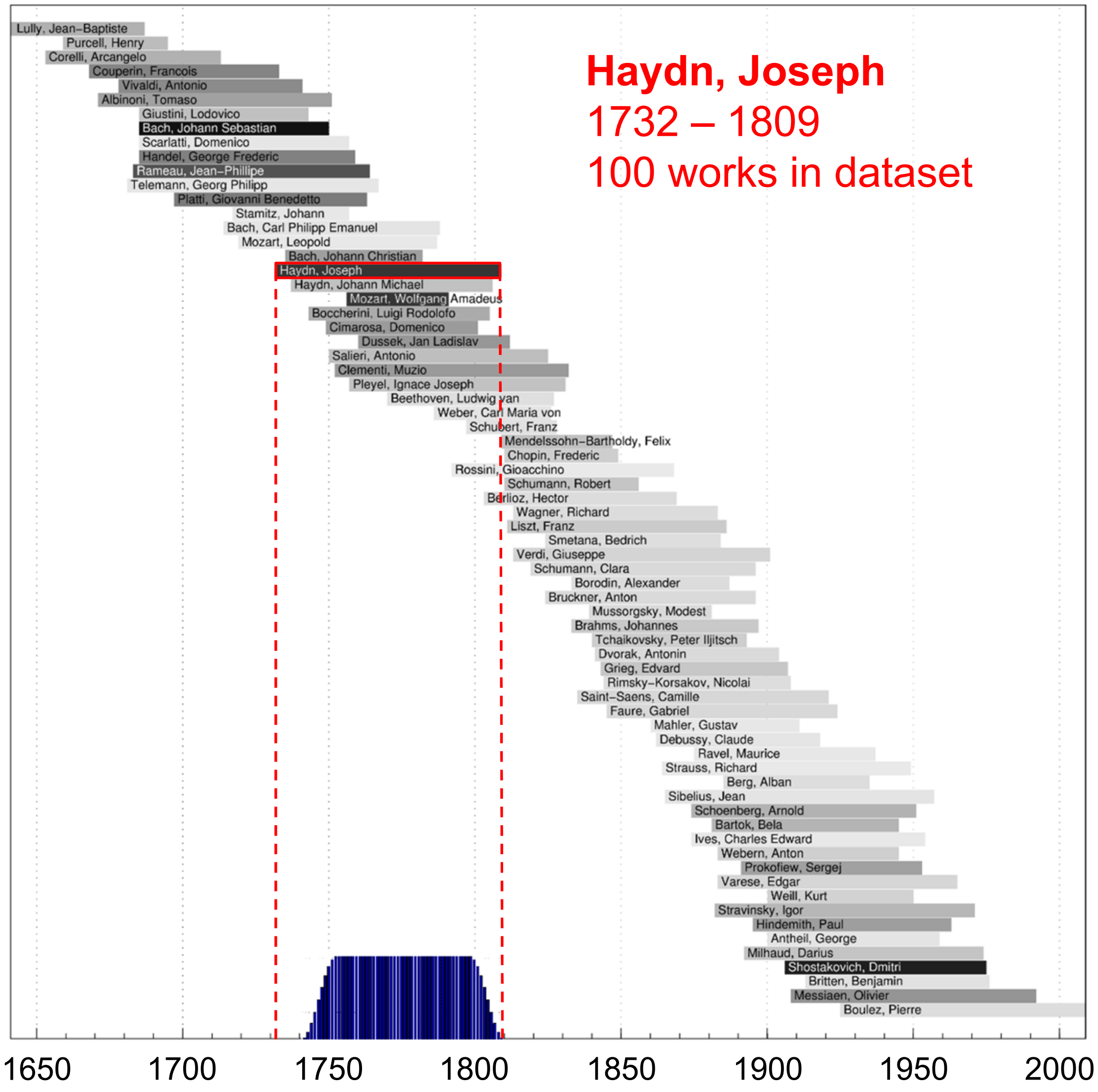
Recall: Tonal Complexity



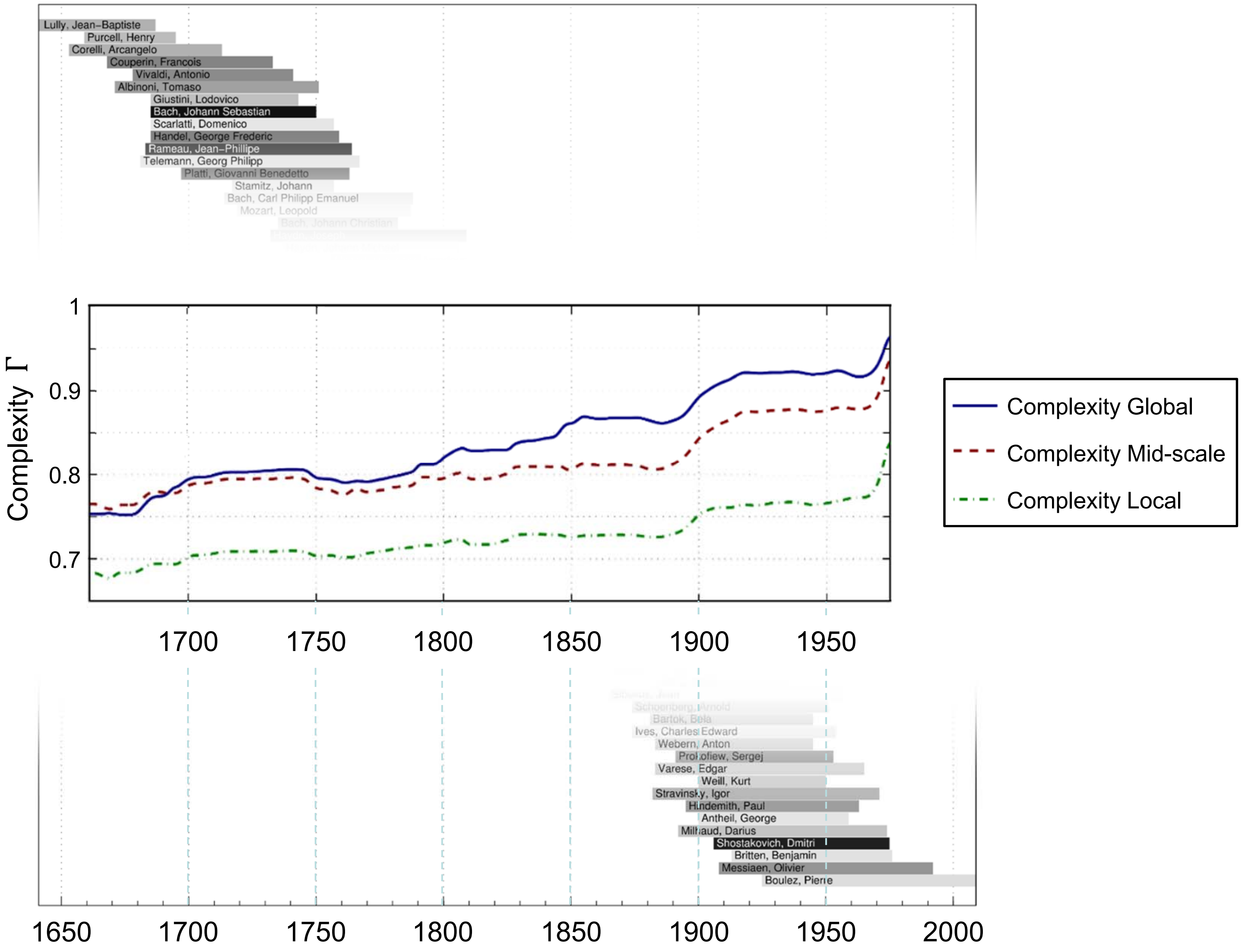
Recall: Tonal Complexity



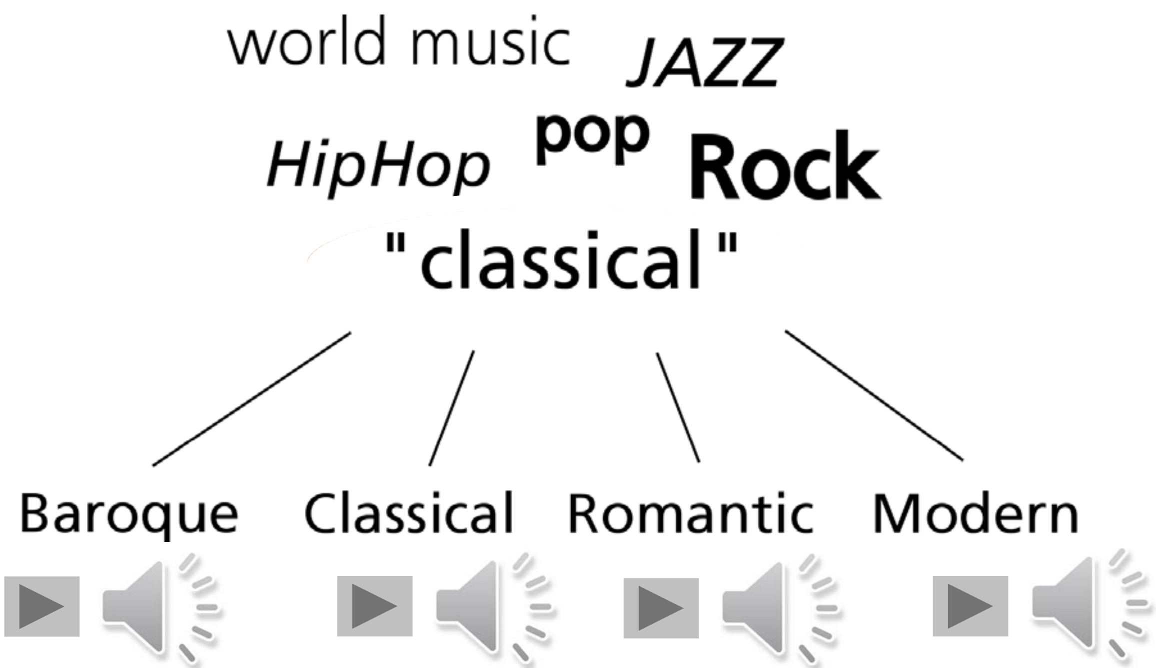
Recall: Tonal Complexity



Recall: Tonal Complexity



Music Genre Classification



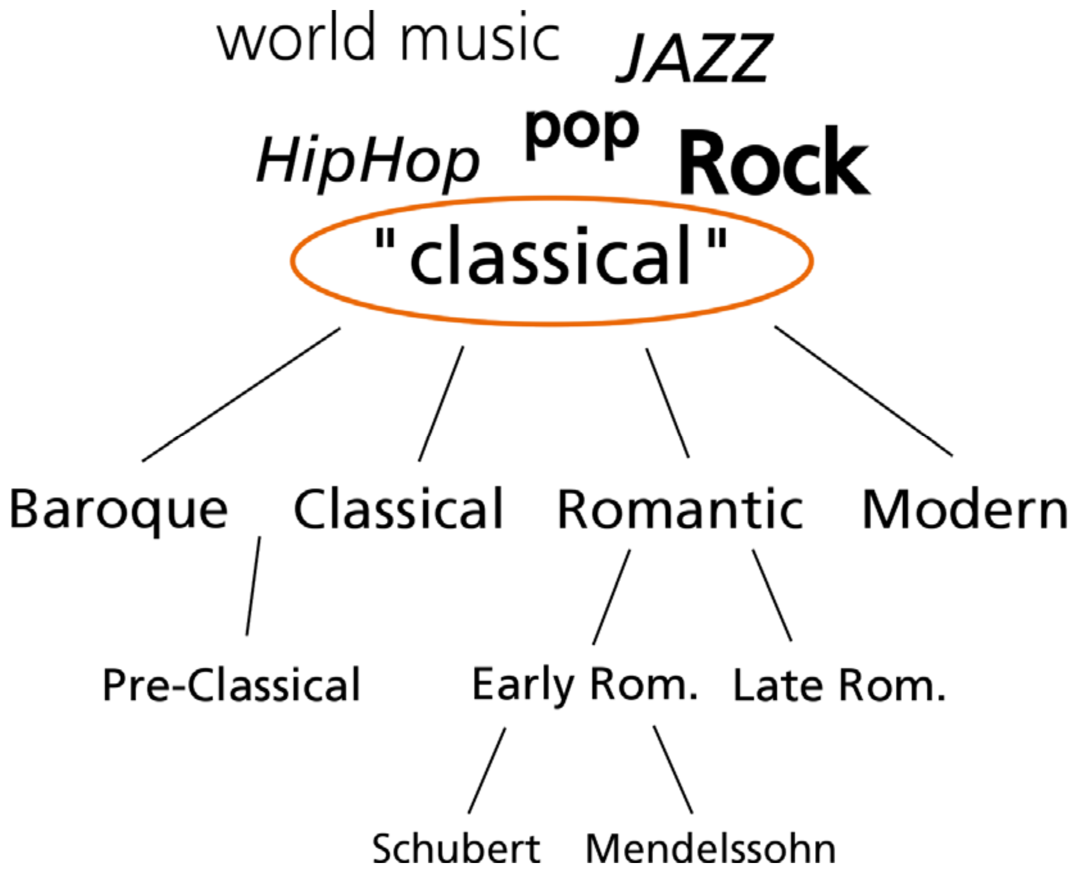
J. S. Bach,
Brandenburg Concerto
No. 2 in F major, I. Allegro,
Cologne Chamber Orch.

L. van Beethoven,
Fidelio, Overture,
Slovak Philharm.

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

A. Webern,
Variations for Orchestra op. 30
Ulster Orchestra

Music Genre Classification



Subgenre Categories:

Period / Era

Sub-era

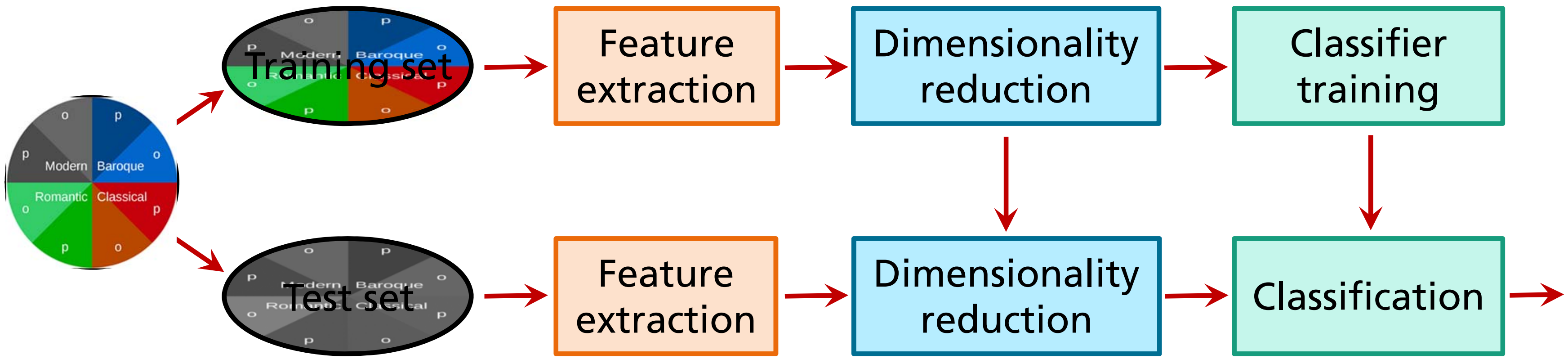
Composer

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

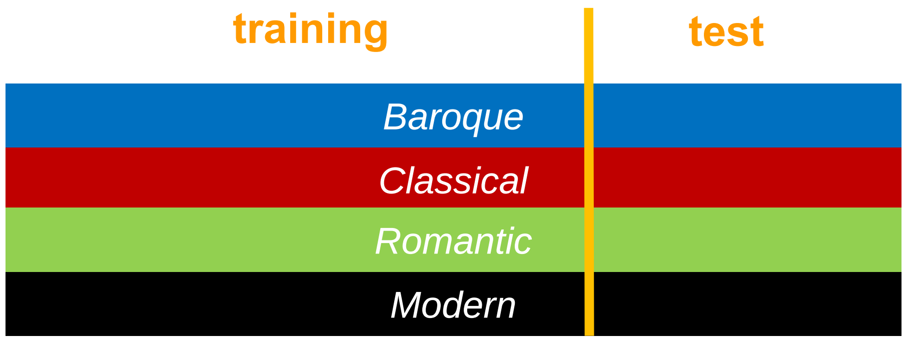


Music Genre Classification

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

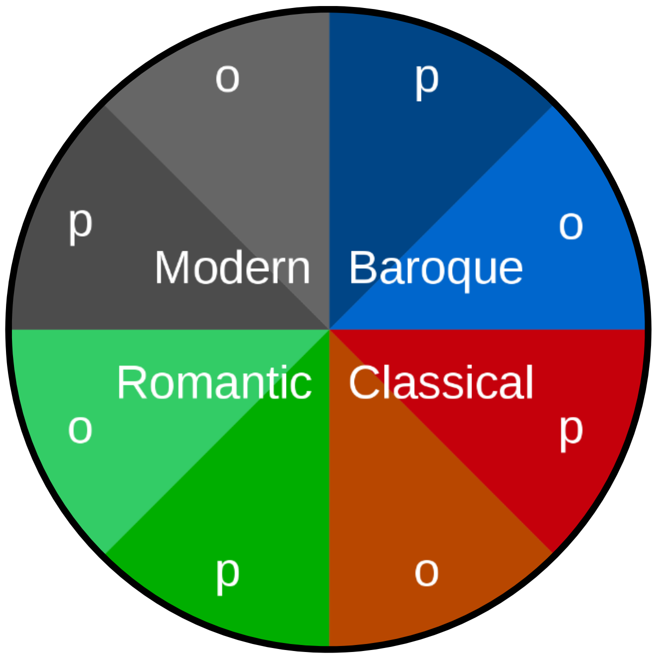
	Fold 1	*Fold 2*	*Fold 3*
Round 1	Training fold	Training fold	Test fold
Round 2	Training fold	Test fold	Training fold
Round 3	Test fold	Training fold	Training fold

- Distribution of classes balanced for all folds

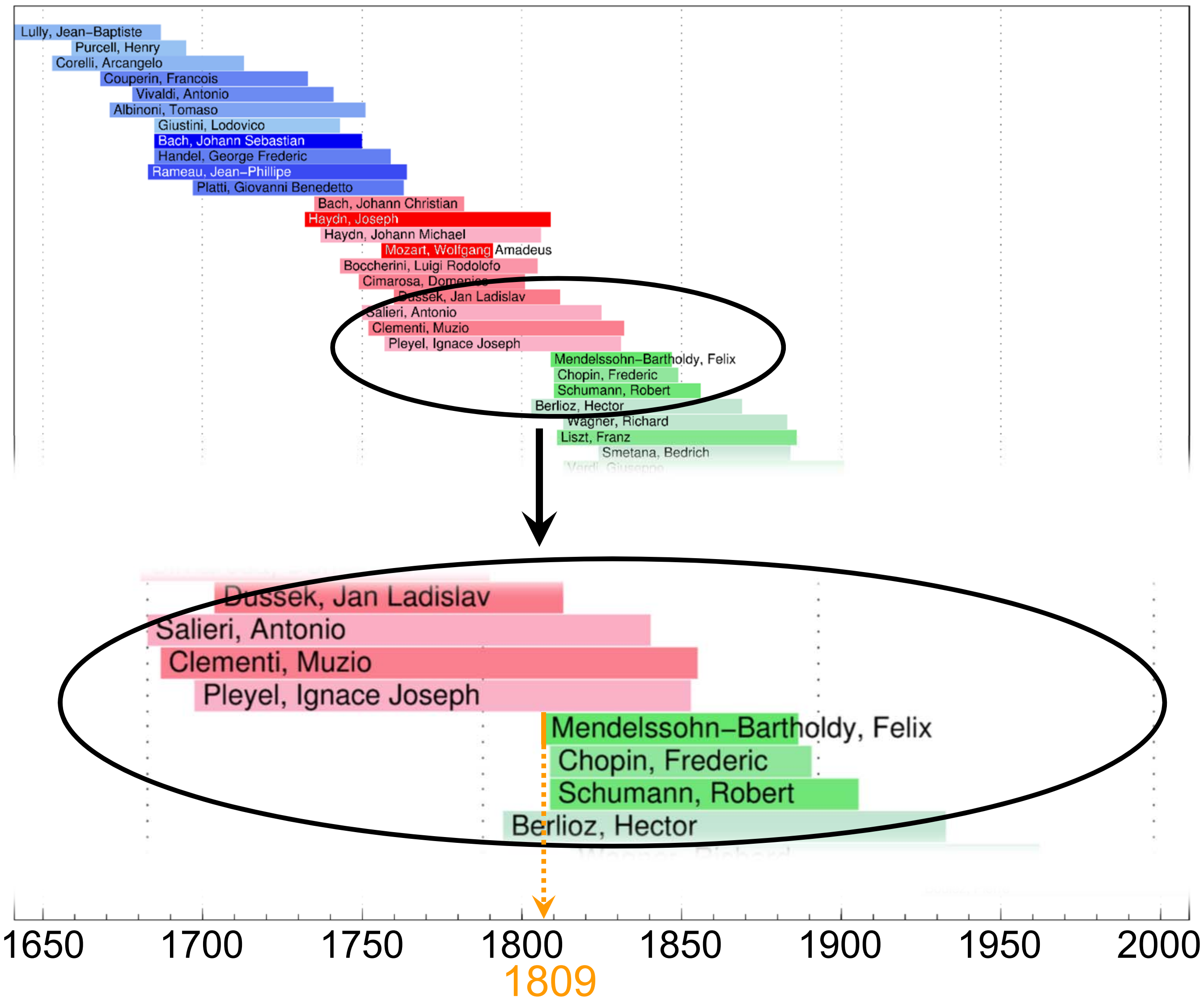


Classification Scenario

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



Classification Scenario

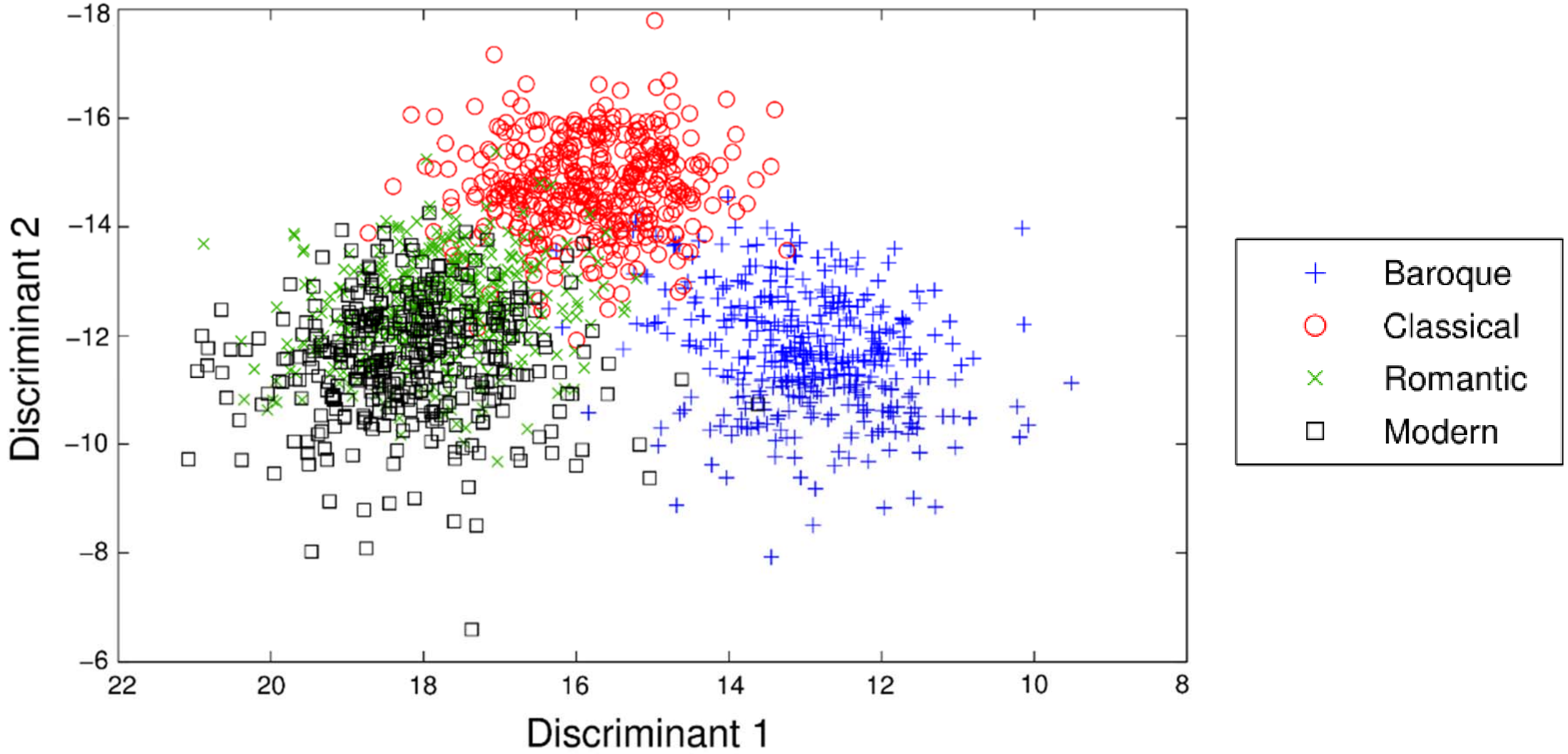


Classification Features

Standard	Dim.	Tonal	Dim.
MFCC	16	Interval cat.	6 x 4
OSC	14	Triad types	4 x 4
ZCR	1	Complexity	7 x 4
ASE	16	Chord progr.	11 x 5
SFM	16		
SCF	16		
SC	16		
LogLoud	12		
NormLoud	12		
Sum	119	Sum	123
Mean & Std	x 2	Mean & Std	x 2
Total	**238**	**Total**	**246**

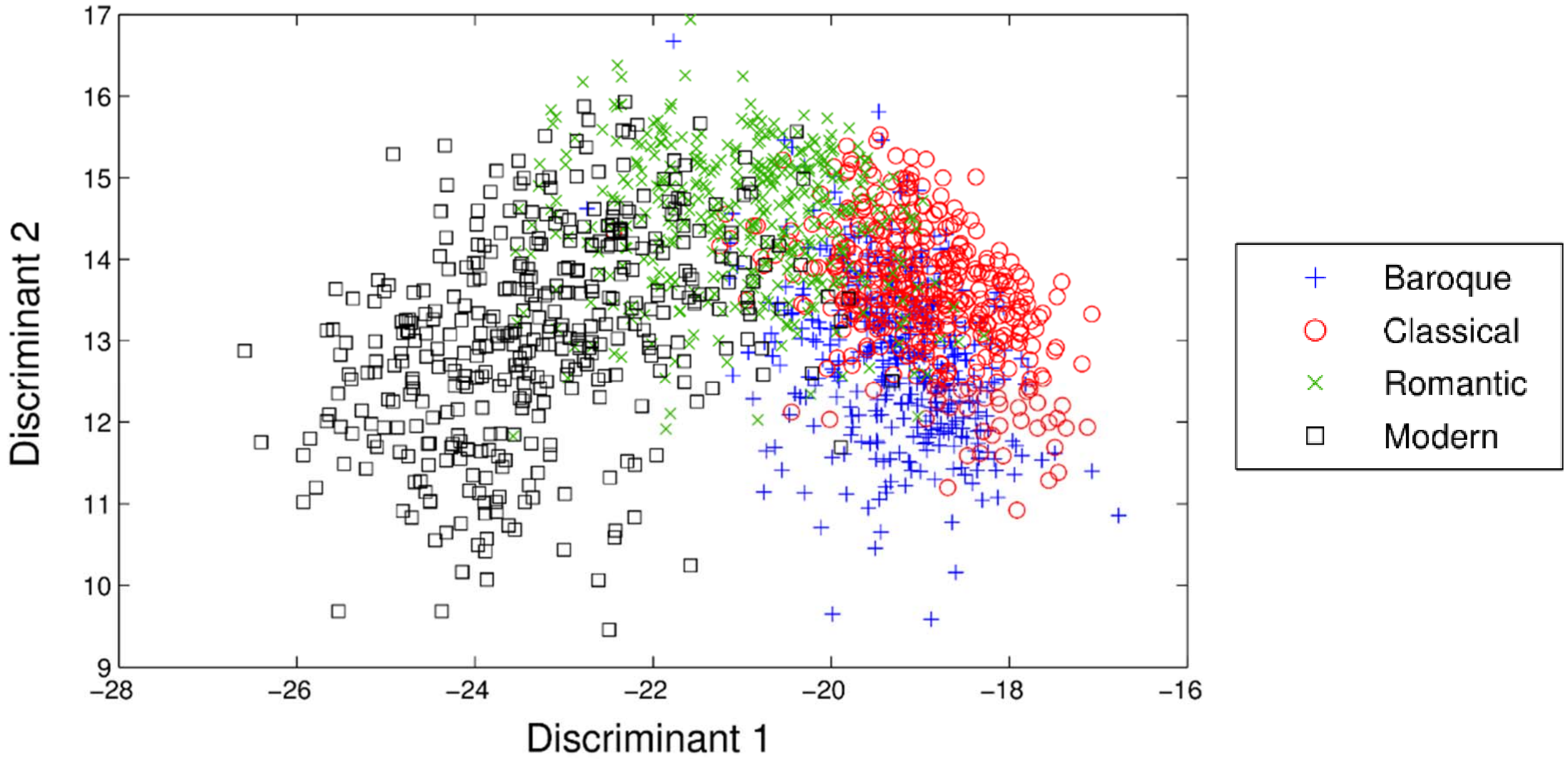
Dimensionality Reduction

- 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, …)



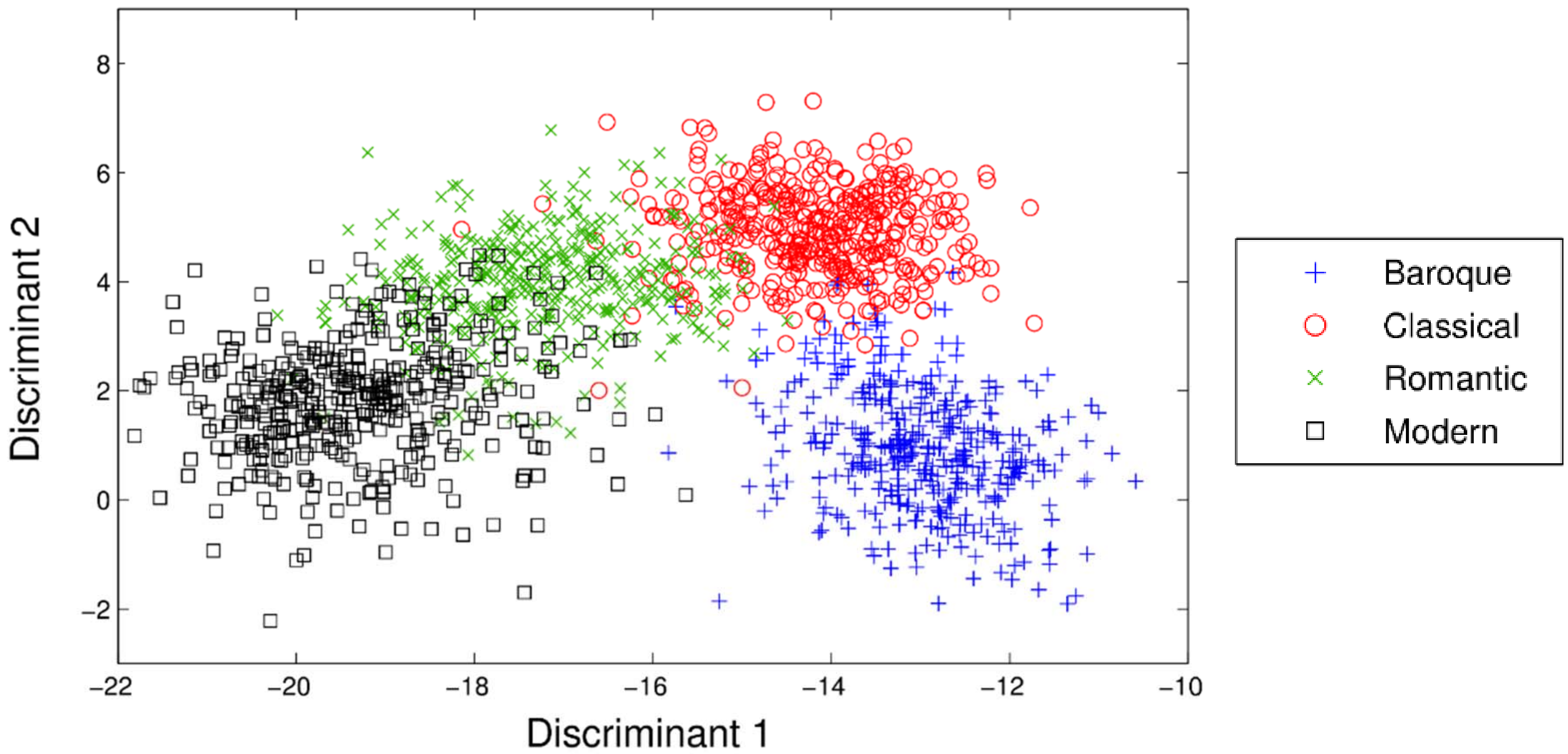
Dimensionality Reduction

- 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

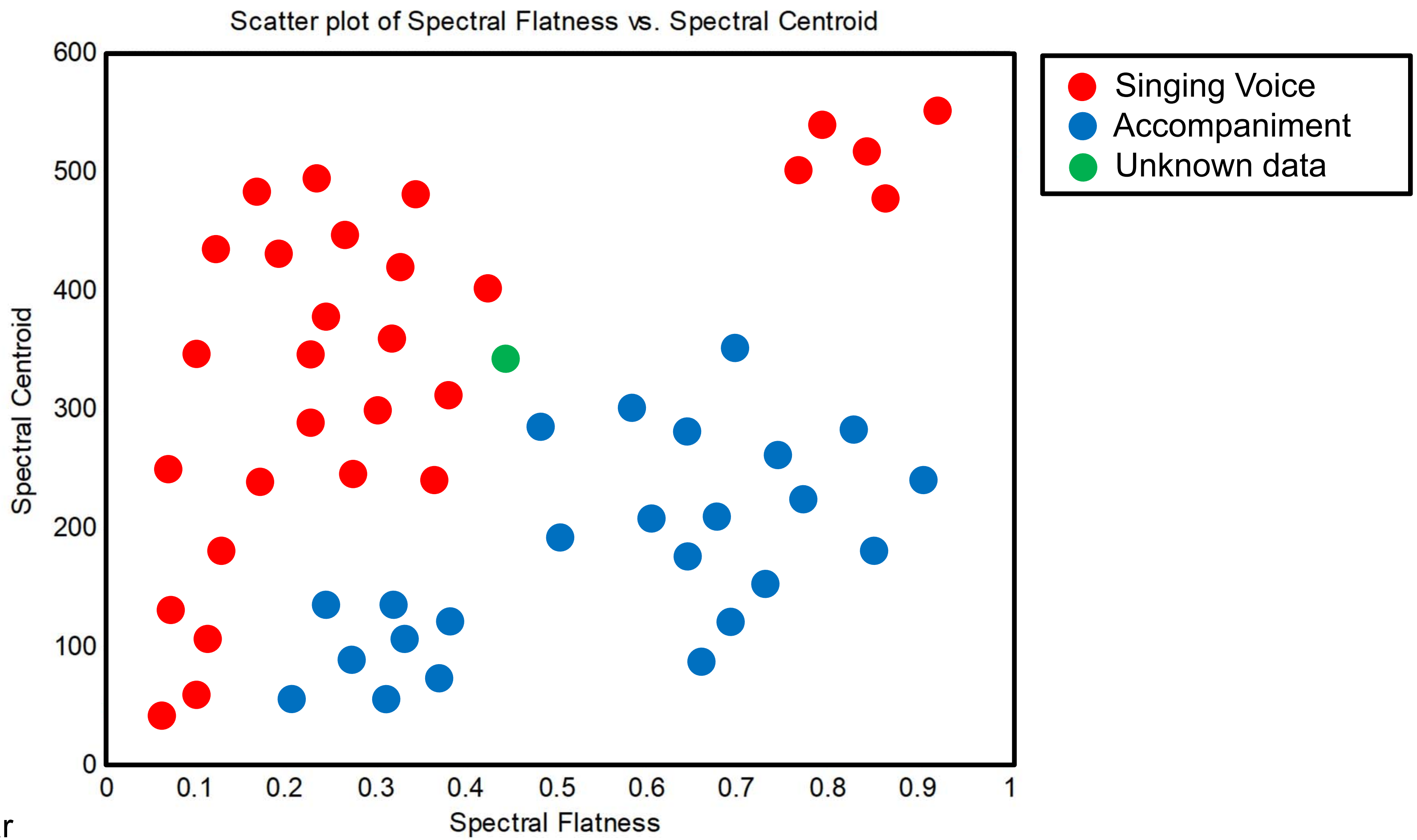
- Reduce feature space to few dimensions

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

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

Classification methods

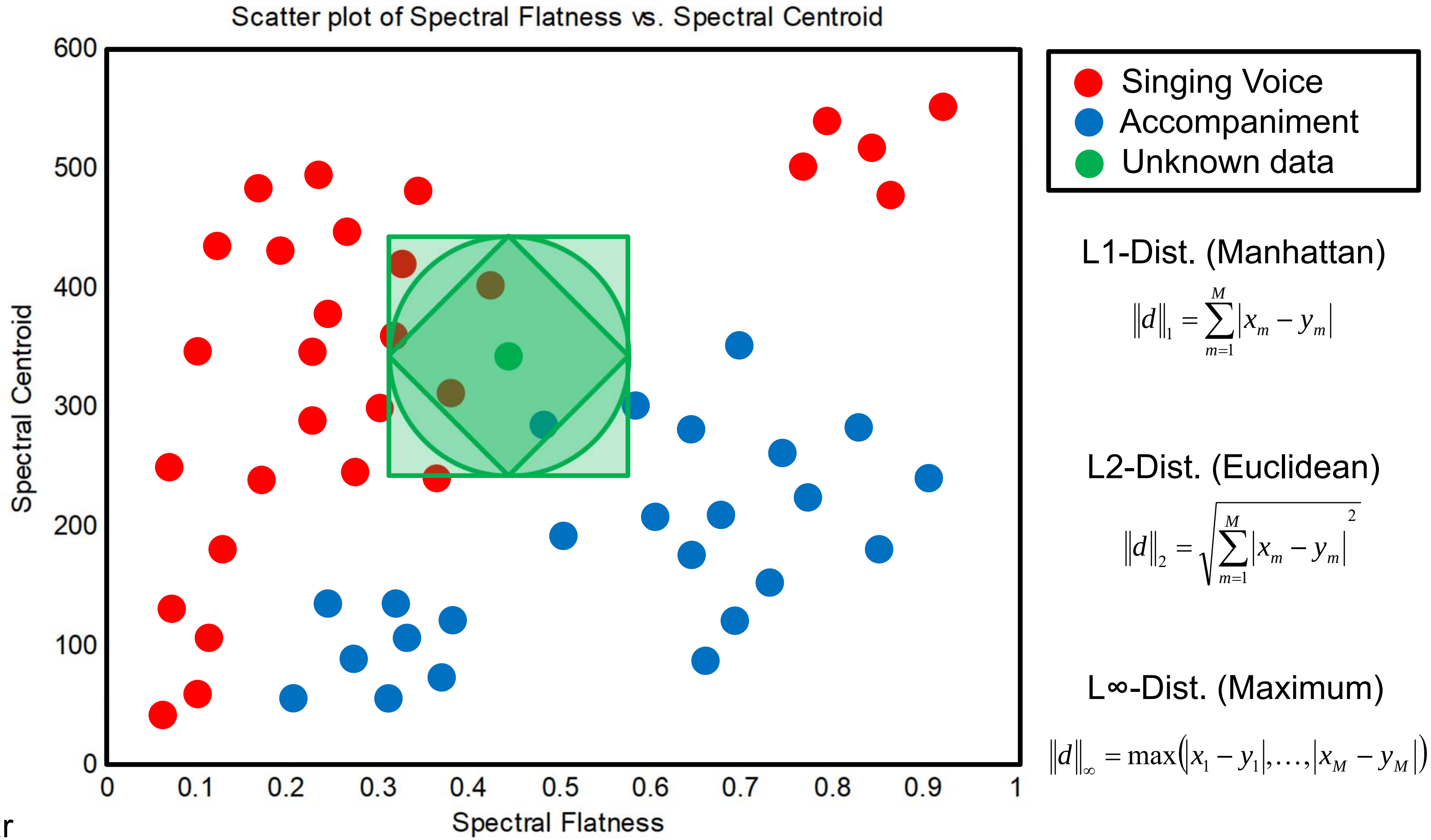
- *k* Nearest Neighbours (kNN)



Slides:
Christian Dittmar

Classification methods

- *k* Nearest Neighbours (kNN)



L1-Dist. (Manhattan)

$$\|d\|_1 = \sum_{m=1}^{M} |x_m - y_m|$$

L2-Dist. (Euclidean)

$$\|d\|_2 = \sqrt{\sum_{m=1}^{M} |x_m - y_m|^2}$$

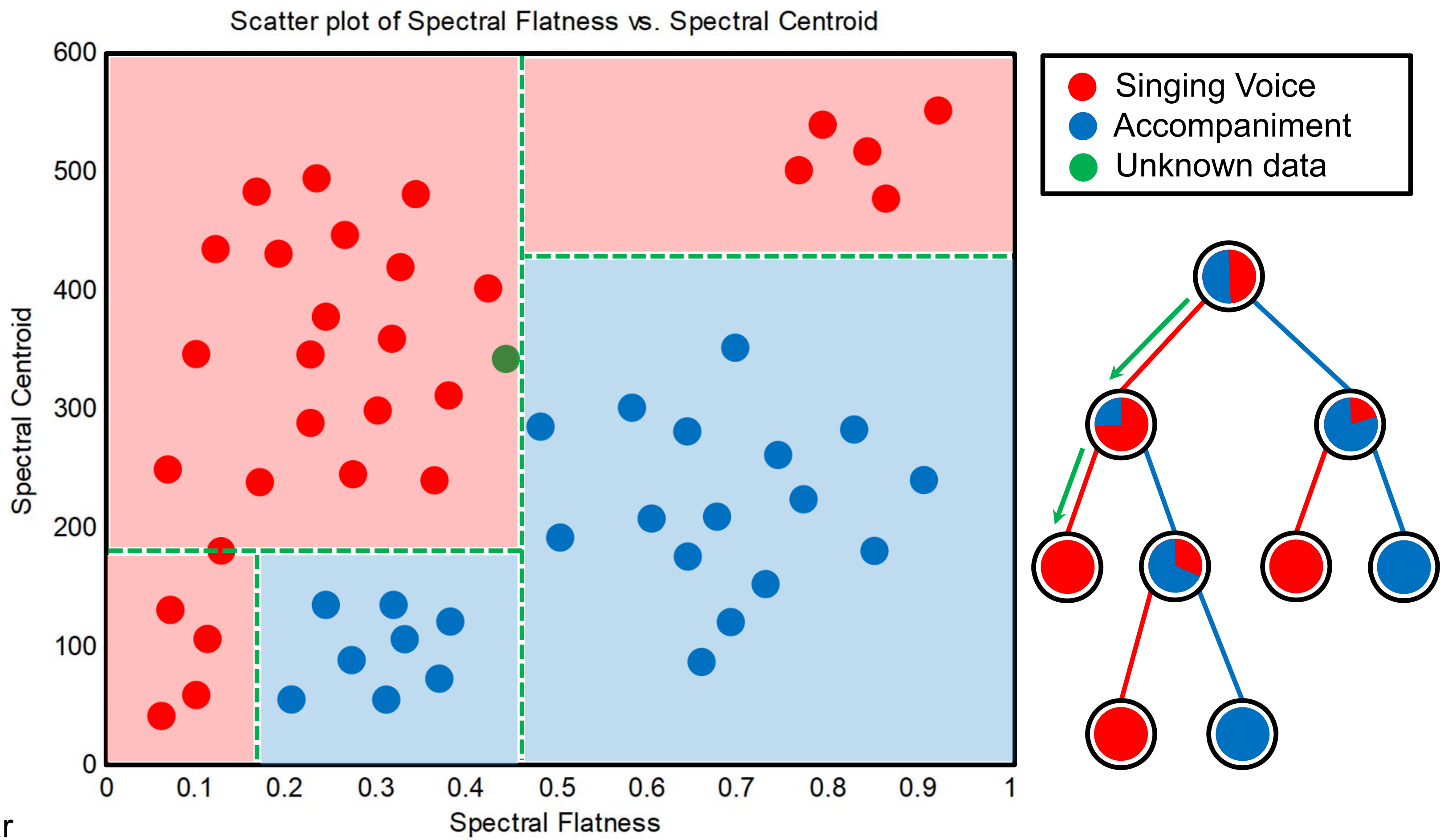
L∞-Dist. (Maximum)

$$\|d\|_\infty = \max(|x_1 - y_1|, \ldots, |x_M - y_M|)$$

Slides:
Christian Dittmar

Classification methods

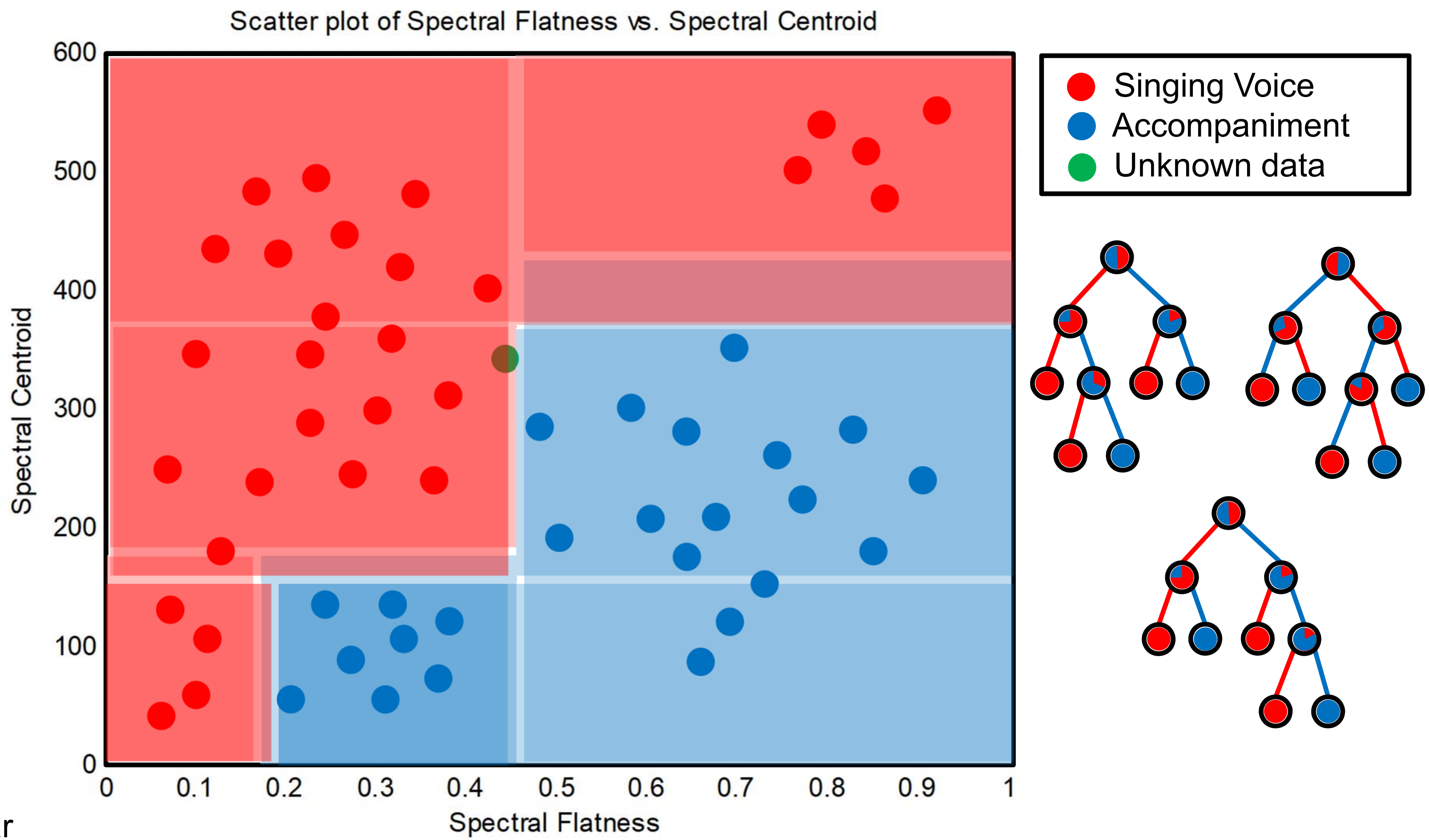
- Decision Trees (DT)



Slides:
Christian Dittmar

Classification methods

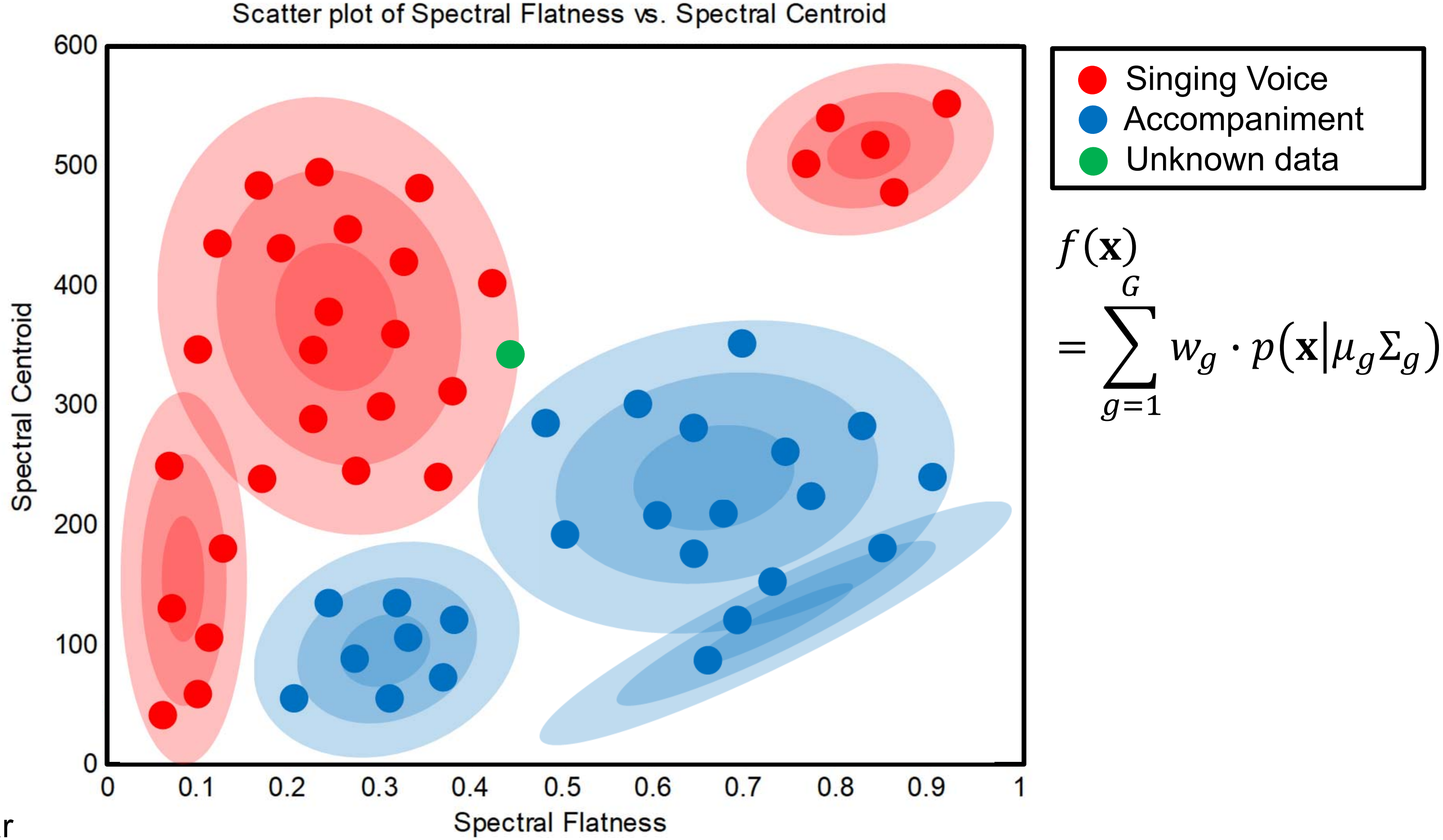
- Random Forests (RF)



Slides:
Christian Dittmar

Classification methods

- Gaussian Mixture Models (GMM)

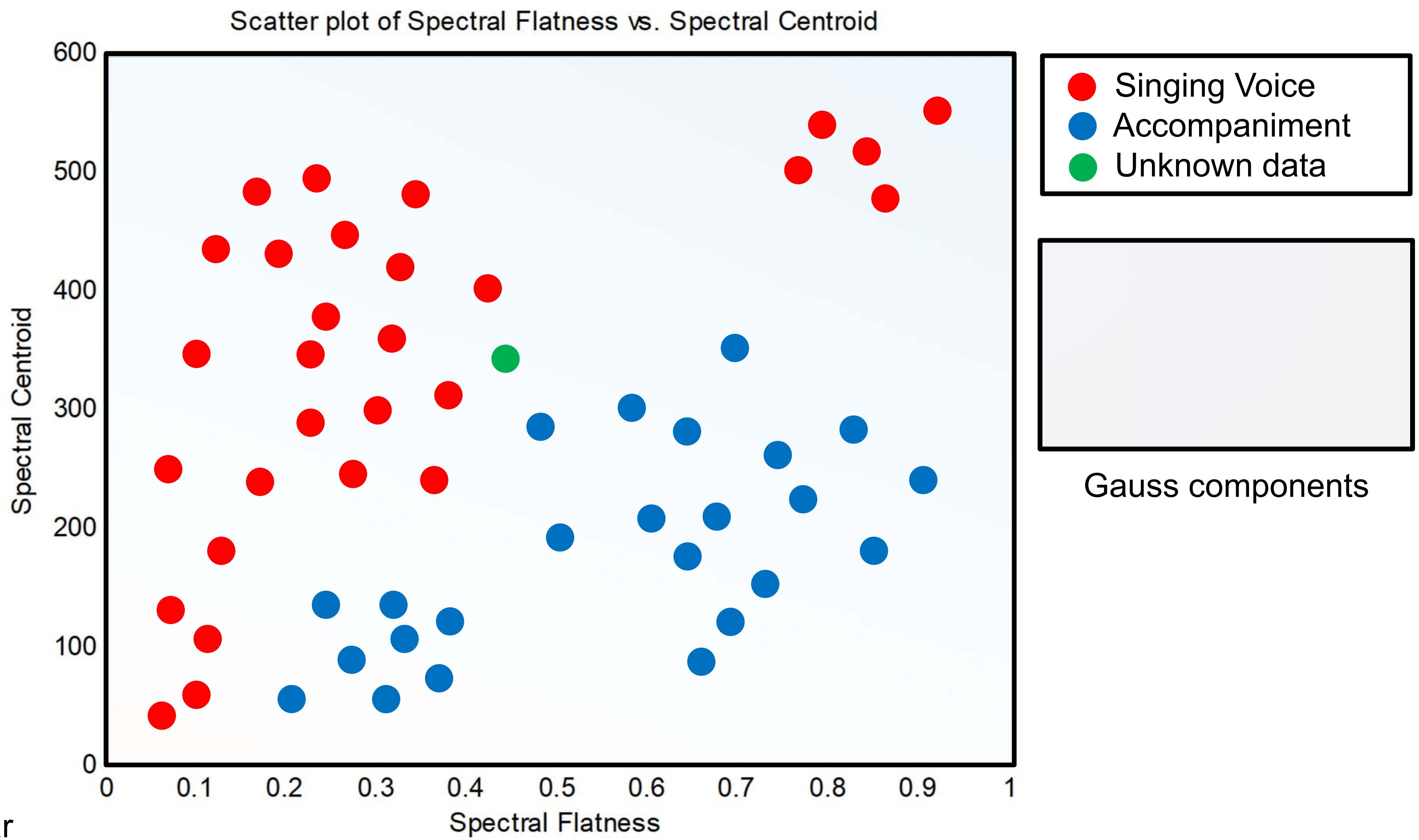


$$f(\mathbf{x}) = \sum_{g=1}^{G} w_g \cdot p\left(\mathbf{x} \middle| \mu_g \Sigma_g\right)$$

Slides:
Christian Dittmar

Classification methods

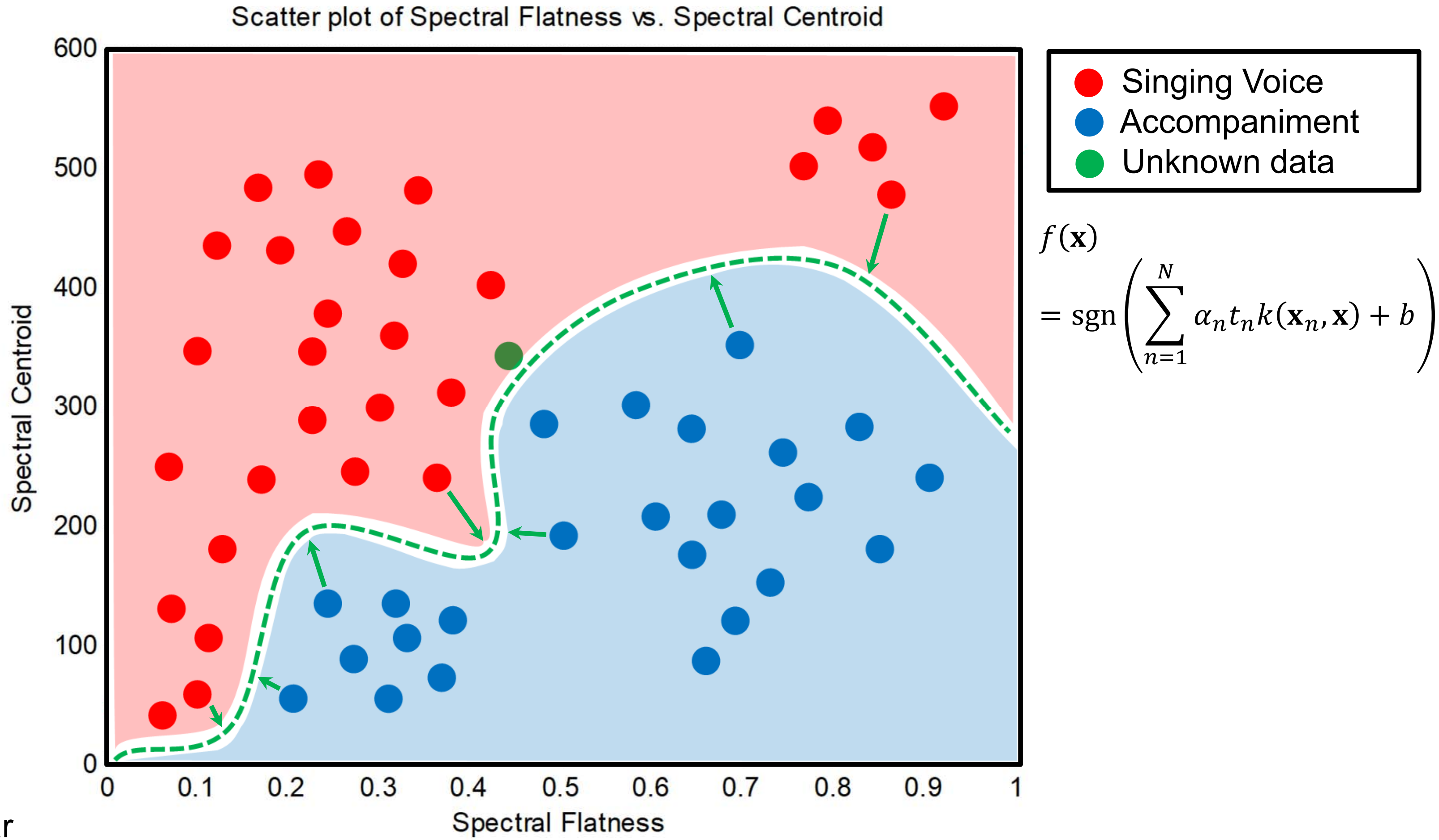
- Gaussian Mixture Models (GMM)



Slides:
Christian Dittmar

Classification methods

- Support Vector Machines (SVM)

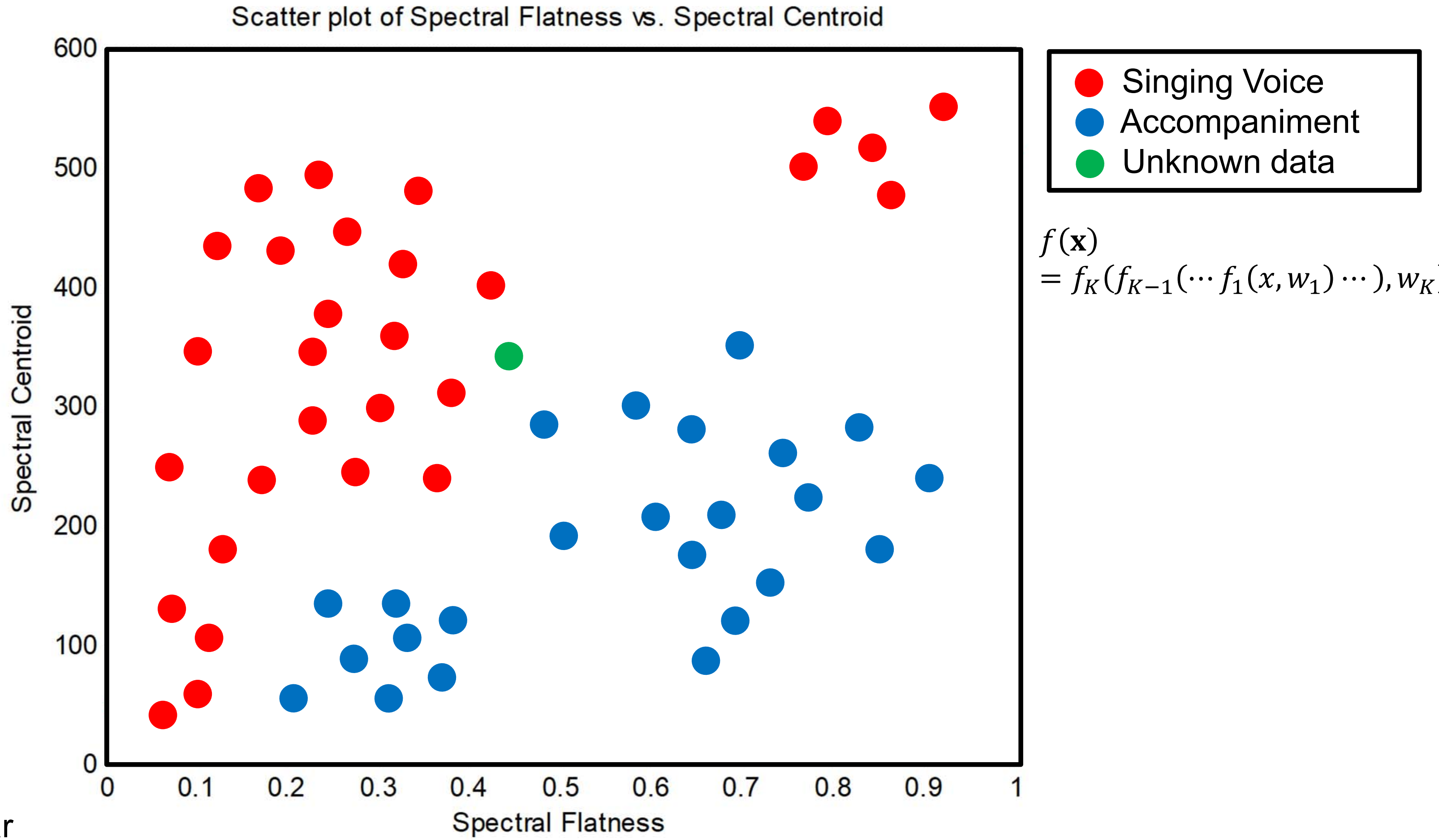


$$f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{n=1}^{N} \alpha_n t_n k(\mathbf{x}_n, \mathbf{x}) + b\right)$$

Slides:
Christian Dittmar

Classification methods

- Deep Neural Networks (DNN)

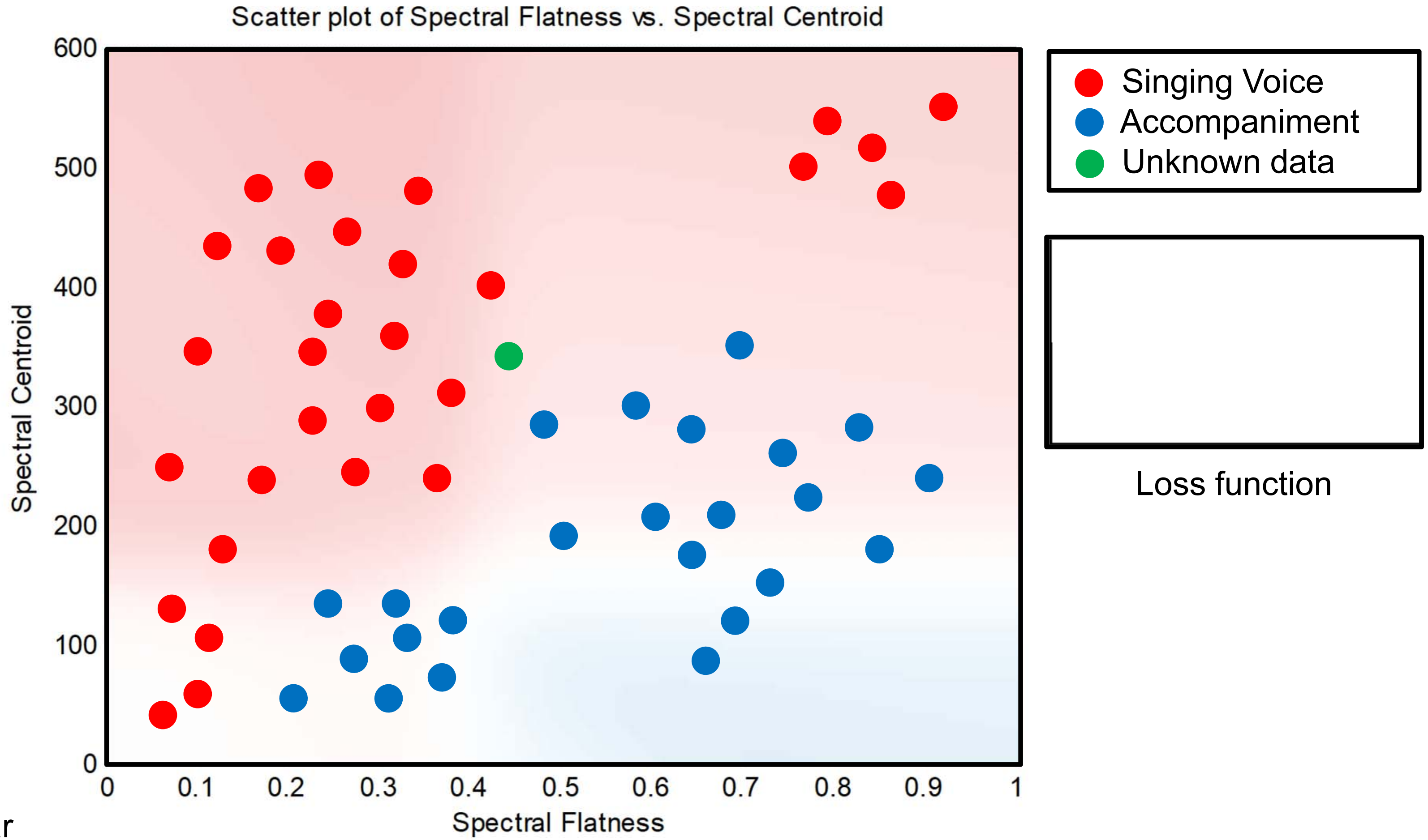


$f(\mathbf{x})$
$= f_K(f_{K-1}(\cdots f_1(x, w_1) \cdots), w_K)$

Slides:
Christian Dittmar

Classification methods

- Deep Neural Networks (DNN)



Slides:
Christian Dittmar

Classification Results

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

	Full Dataset	**Piano**	**Orchestra**
Standard features	87 %	88 %	85 %
Tonal features	84 %	84 %	86 %
Combined	**92 %**	**86 %**	**80 %**

Weiss / Mauch / Dixon, *Timbre-Invariant Audio Features for Style Analysis of Classical Music*, ICMC / SMC 2014

Classification Results

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

	Full Dataset	**Piano**	**Orchestra**
Standard features	87 %	88 %	85 %
Tonal features	84 %	84 %	86 %
Combined	**92 %**	**86 %**	**80 %**

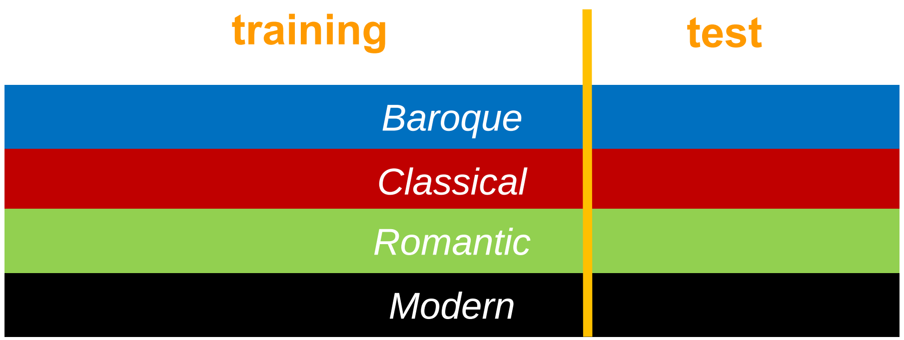
Overfitting???

Weiss / Mauch / Dixon, *Timbre-Invariant Audio Features for Style Analysis of Classical Music*, ICMC / SMC 2014

Classification Results

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

	Full Dataset	**Piano**	**Orchestra**
Standard features	87 %	88 %	85 %
Tonal features	84 %	84 %	86 %
Combined	**92 %**	**86 %**	**80 %**



Flexer, *A Closer Look on Artist Filters for Musical Genre Classification*, ISMIR 2007

Classification Results

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

	Full Dataset	**Piano**	**Orchestra**
Standard features	87 %	88 %	85 %
Tonal features	84 %	84 %	86 %
Combined	**92 %**	**86 %**	**80 %**

- **Using composer filter**

	Full Dataset	**Piano**	**Orchestra**
Standard features	54 %	36 %	70 %
Tonal features	73 %	70 %	78 %
Combined	**68 %**	**44 %**	**68 %**

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**

	Full Dataset	Piano	Orchestra
Standard features	87 %	88 %	85 %
Tonal features	84 %	84 %	86 %
Combined	**92 %**	**86 %**	**80 %**

- **Using composer filter**

	Full Dataset	Piano	Orchestra
Standard features	54 %	36 %	70 %
Tonal features	73 %	70 %	78 %
Combined	**68 %**	**44 %**	**68 %**

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 %**

Era (correct) \ Era (classified)	Baroque	Classical	Romantic	Modern
Baroque	65.2	23.2	10.9	0.6
Classical	17.0	74.9	8.1	0.0
Romantic	6.5	5.0	77.7	10.8
Modern	1.7	0.9	16.8	80.6

Classification Results: Error Examples

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

Class	*Composer*	*Piece*	*Classified*
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in E♭ minor BWV 853	Romantic
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in F major BWV 856	Romantic
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in A minor BWV 865	Romantic
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in B♭ major BWV 866	Romantic
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in B♭ minor BWV 867	Romantic
Baroque	Bach, J. S.	English Suite No. 3 in G minor BWV 808, Sarabande	Romantic
Baroque	Bach, J. S.	Brandenburg Conc. No. 1 in F major BWV 1046, Adagio	Romantic
Baroque	Bach, J. S.	Overture No. 2 in B minor BWV 1067, Badinerie	Romantic
Baroque	Bach, J. S.	Overture No. 3 in D major BWV 1068, Gigue	Romantic
Baroque	Couperin, F.	27 Ordres, Huitième ordre, IX. Rondeau passacaille	Romantic
Baroque	Corelli, A.	Concerto grosso op. 6 No. 2, III. Grave – Andante largo	Romantic
Baroque	Lully, J.-B.	Ballet de Xerces LWV 12, Gavotte en rondeau	Romantic
Baroque	Purcell, H.	Opera "Dido and Aeneas" Z. 626, Overture	Romantic
Baroque	Vivaldi, A.	"The Four Seasons," RV 293 "Autumn," Adagio molto	Romantic
Romantic	Schumann, R.	Kinderszenen op. 15, "Haschemann"	Baroque
Romantic	Grieg, E.	Holberg suite op. 40, Gavotte	Baroque
Romantic	Mendelssohn, F.	Symphony No. 4 in A major, IV. Saltarello, presto	Baroque
Modern	Shostakovich, D.	Preludes & Fugues op. 87 Fugue No. 1 in C major	Baroque
Modern	Shostakovich, D.	Preludes & Fugues op. 87 Fugue No. 5 in D major	Baroque

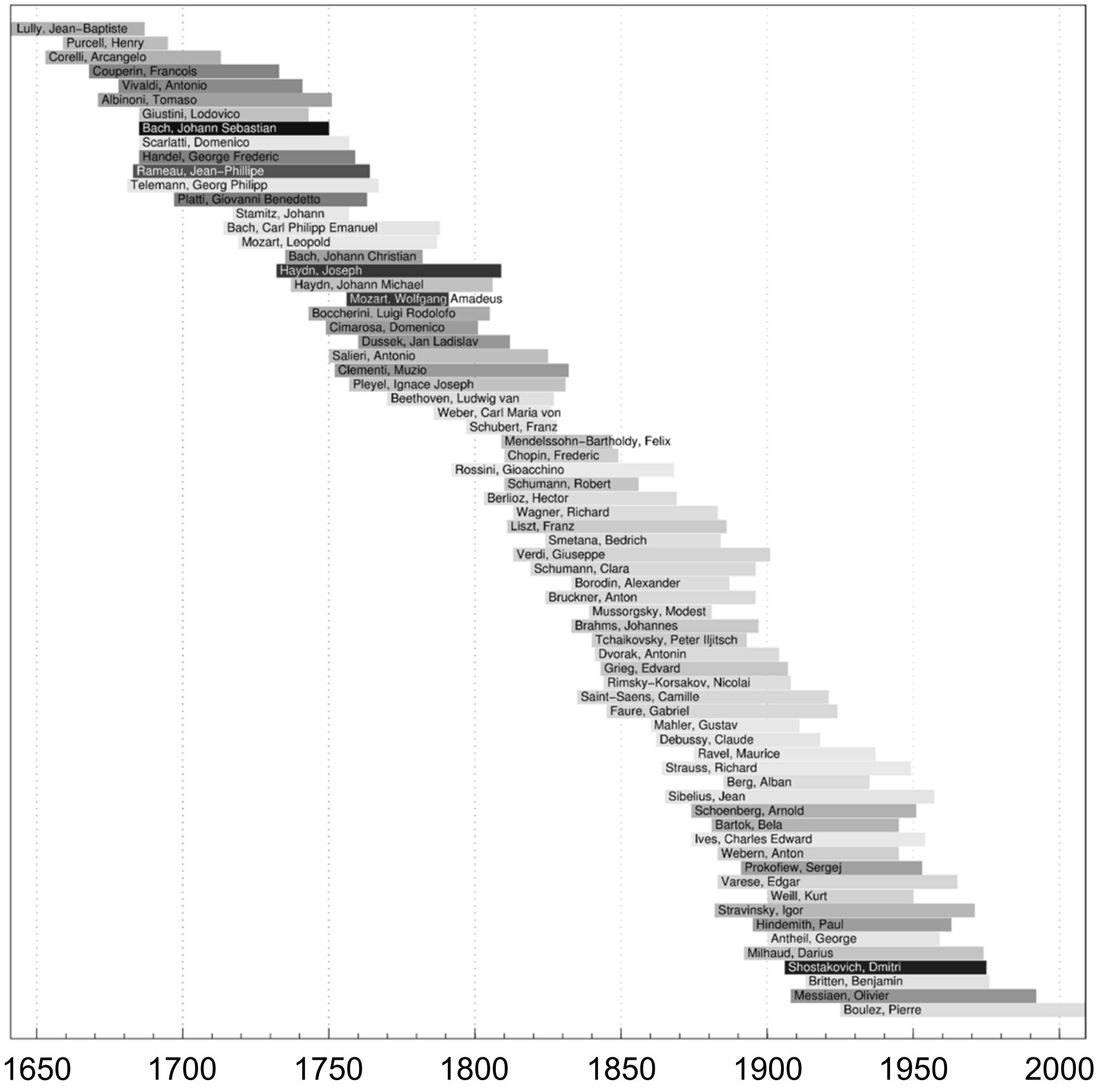
Classification Results – Summary

- Different types of tonal features
- Combination of time scales
- Classifiers (SVM, Random Forest)
- State-of-the-art
  - Few studies on audio
  - Good separation of tonal-vs.-atonal (**91 %**):

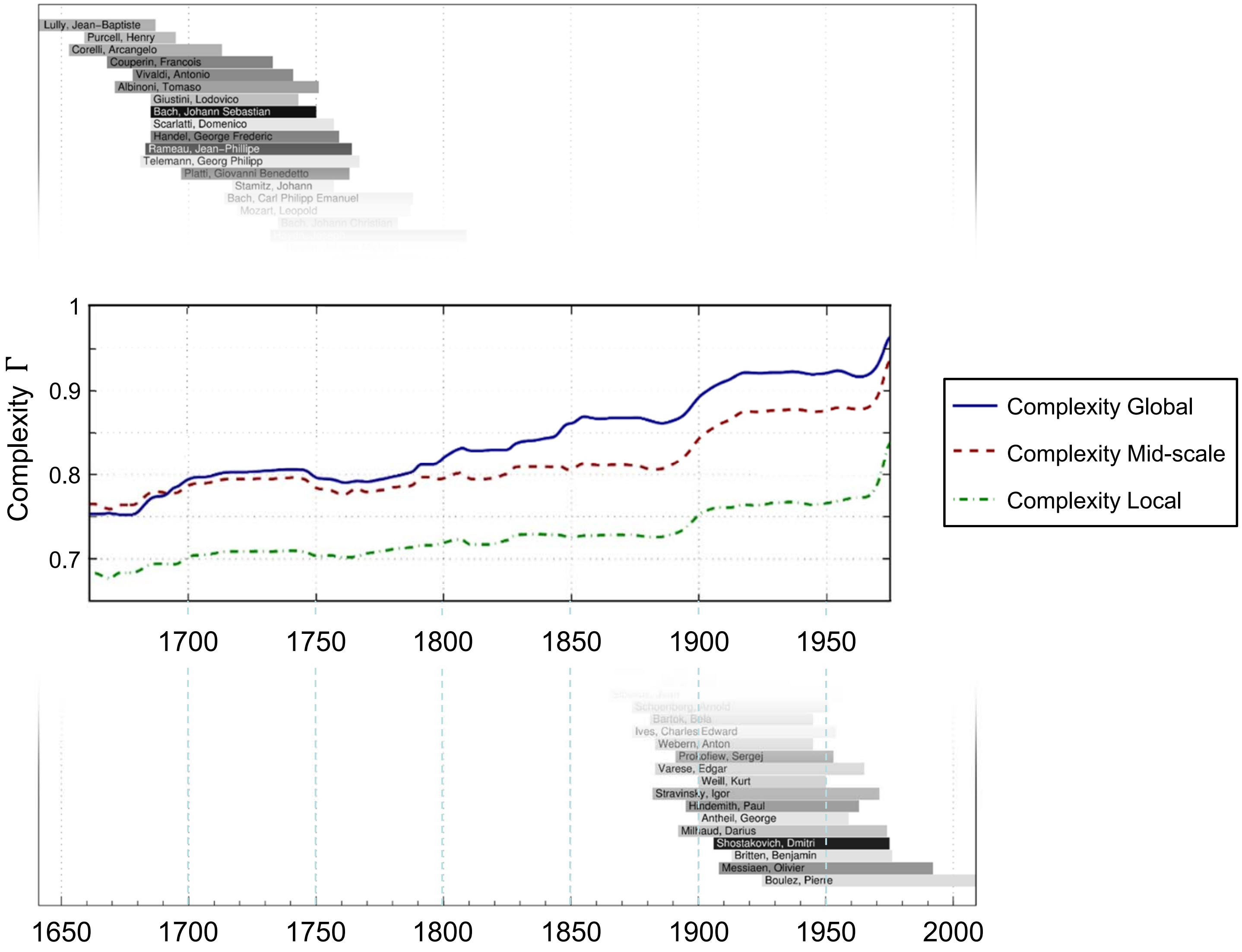
    Izmirli, *Tonal-Atonal Classification of Music Audio Using Diffusion Maps*, ISMIR 2009
- Composer Identification
  - Up to **78 %** for 11 composers

    Hamel, *Pooled Features Classification*, MIREX 2011
  - Dataset balanced?

Musical Style Analysis

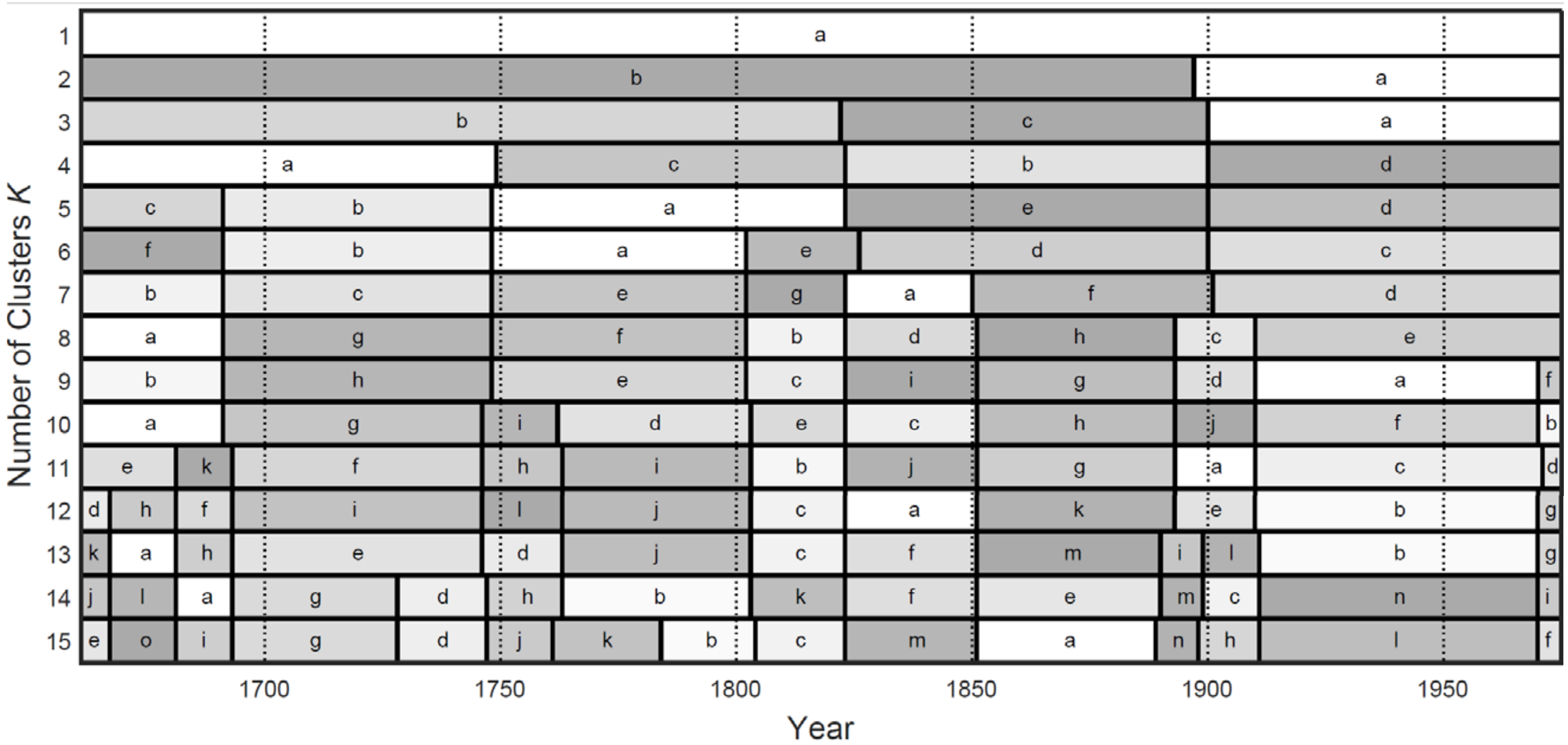


Musical Style Analysis – Complexity



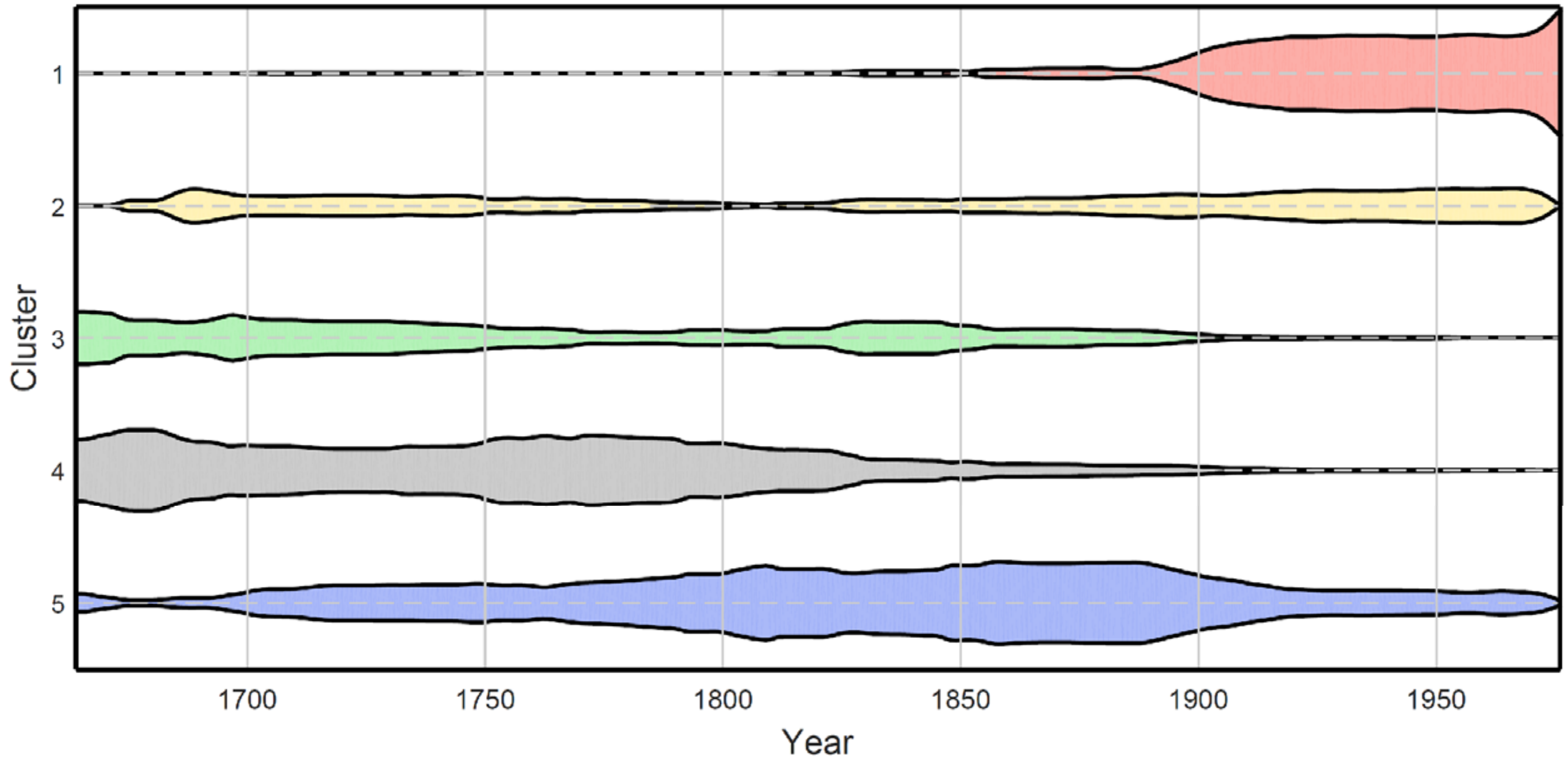
Clustering: Years

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



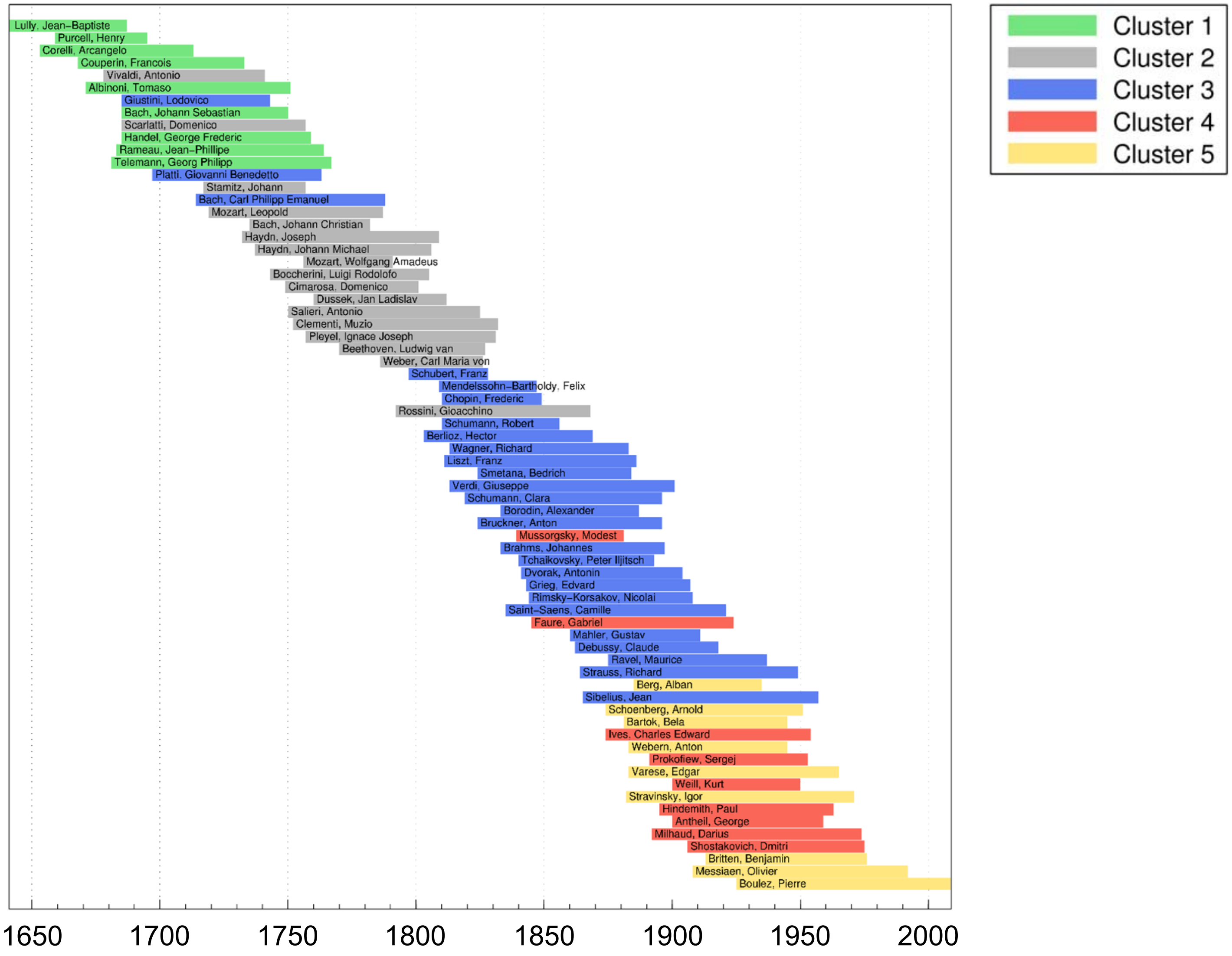
Clustering: Pieces

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

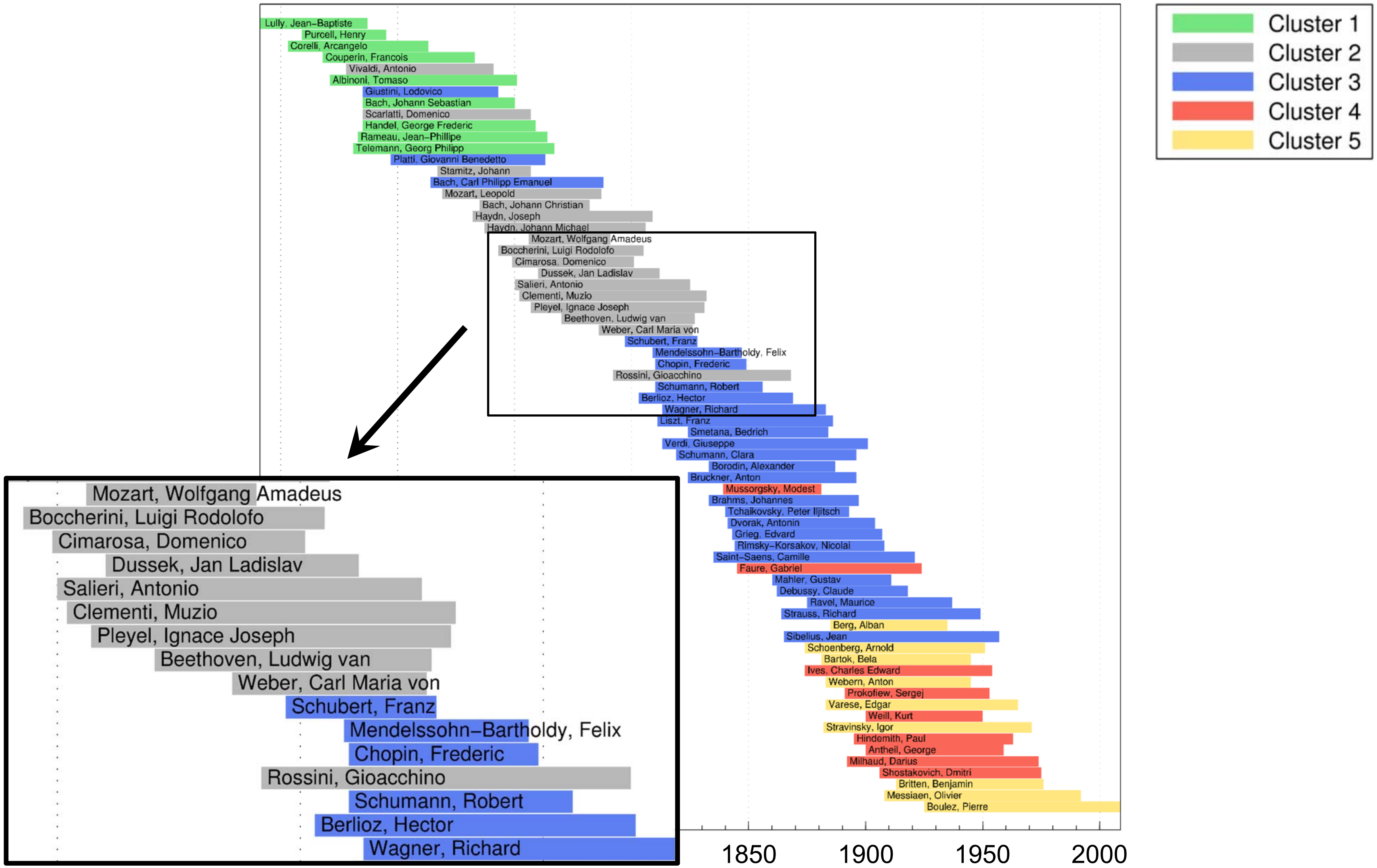


Weiss / Mauch / Dixon / Müller,
Investigating Style Evolution of Western Classical Music: A Computational Approach
Musicae Scientiae 2018

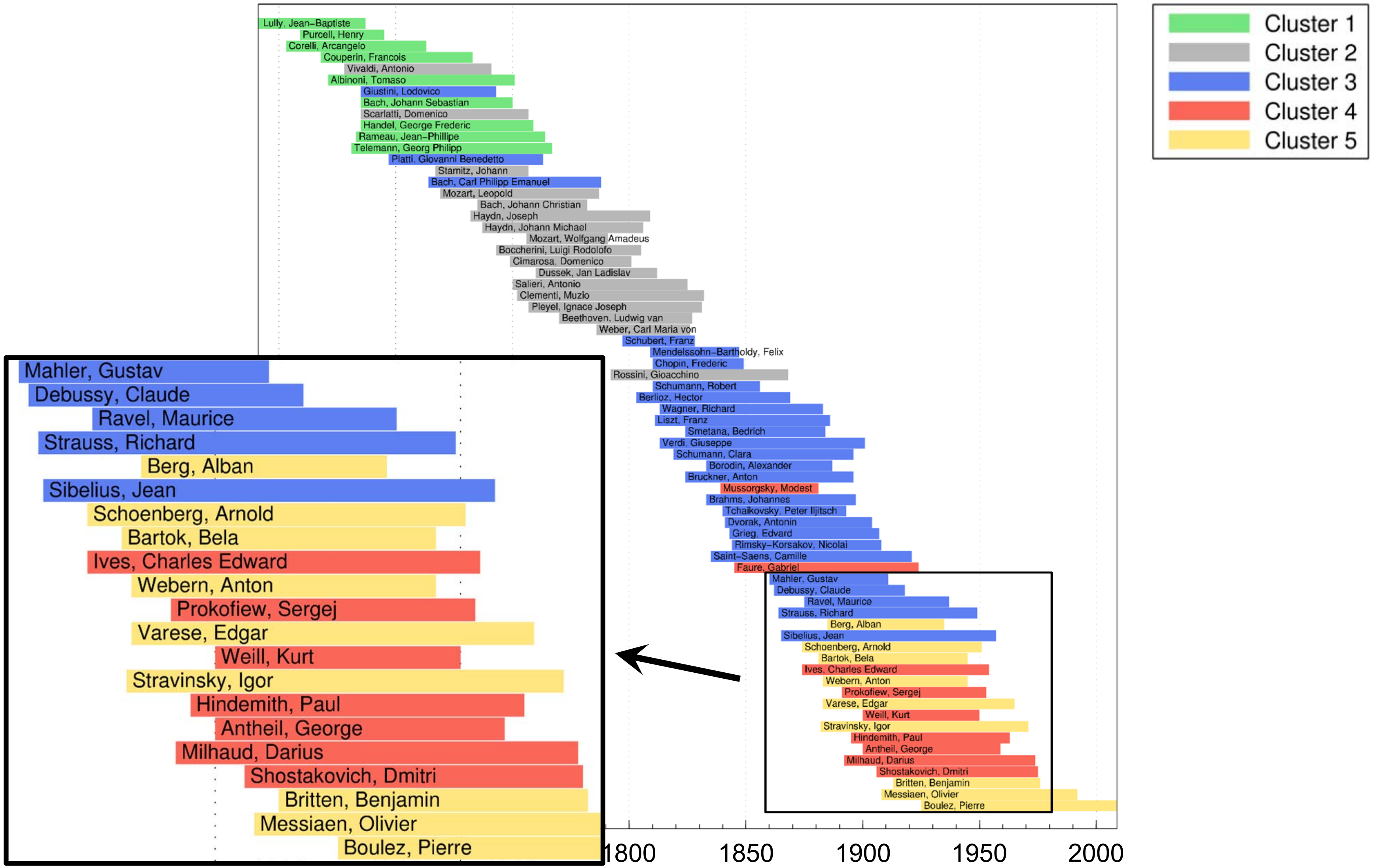
Clustering: Composers



Clustering: Composers



Clustering: Composers



Clustering: Composers