INTERNATIONAL AUDIO LABORATORIES ERLANGEN

AUDIO LABS

Hochschule für Musik Karlsruhe

Blockvorlesung

Advanced Audio-Based Music Processing

7. Style Classification

Christof Weiß, Frank Zalkow, Meinard Müller

International Audio Laboratories Erlangen

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

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

University of Music Hochschule für Musik Karlsruhe

Fraunhofer IIS

Dissertation: Tonality-Based Style Analysis

Christof Weiß
Computational Methods for Tonality-Based Style Analysis of Classical Music Audio Recordings
PhD thesis, Ilmenau University of Technology, 2017
https://www.db-thueringen.de/receive/dbt_mods_00032890

Chapter 8: Subgenre Classification for Western Classical Music

Style Classification

Overview

Machine Learning pipeline:

- Feature extraction
- Classification

Style Classification
Overview

Machine Learning pipeline:

- Feature extraction
- Classification

Motivation

Music consumption 2000

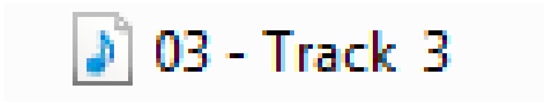
Name	Typ	Größe	Ordner
01 - Track 1	MP3-Audioformat	6.600 KB	Afterworld - Dark ...
01 - Track 1	MP3-Audioformat	6.613 KB	Eidolon - Hallowe...
01 -Track1	MP3-Audioformat	5.214 KB	Symphony X - Da...
01__TRACK_1	MP3-Audioformat	8.705 KB	Pain Of Salvation ...
02 - Track 2	MP3-Audioformat	3.222 KB	Symbyosis - Crisis...
02 - Track 2	MP3-Audioformat	6.517 KB	Afterworld - Dark ...
02 - Track 2	MP3-Audioformat	7.285 KB	Eidolon - Hallowe...
02 -Track2	MP3-Audioformat	8.774 KB	Symphony X - Da...
02__TRACK_2	MP3-Audioformat	1.938 KB	Pain Of Salvation ...
03 - Track 3	MP3-Audioformat	8.221 KB	Symbyosis - Crisis...
03 - Track 3	MP3-Audioformat	7.077 KB	Afterworld - Dark ...
03 - Track 3	MP3-Audioformat	7.067 KB	Eidolon - Hallowe...
03 -Track3	MP3-Audioformat	7.924 KB	Symphony X - Da...
03__TRACK_3	MP3-Audioformat	9.195 KB	Pain Of Salvation ...
03_nightwish-nightquest_(japanese_bonustrack)-amrc	MP3-Audioformat	5.934 KB	Nightwish - The_...
04 -- The way it used to be (non lp track)	MP3-Audioformat	9.206 KB	dreamtheater - 19...
04 - Track 4	MP3-Audioformat	4.716 KB	Symbyosis - Crisis...
04 - Track 4	MP3-Audioformat	7.264 KB	Afterworld - Dark ...
04 - Track 4	MP3-Audioformat	7.849 KB	Eidolon - Hallowe...
04 -Track4	MP3-Audioformat	5.722 KB	Symphony X - Da...
04__TRACK_4	MP3-Audioformat	12.835 KB	Pain Of Salvation ...
05 - Track 5	MP3-Audioformat	6.687 KB	Symbyosis - Crisis...
05 - Track 5	MP3-Audioformat	7.019 KB	Afterworld - Dark ...
05 - Track 5	MP3-Audioformat	7.647 KB	Eidolon - Hallowe...
05 -Track5	MP3-Audioformat	6.150 KB	Symphony X - Da...

Motivation

Music consumption 2000

I want to play that!
Automatic transcription
Melody extraction
Chord recognition
Key determination

Which content?
Speech / music / noise?
Audio classification



What is representative?
Audio Thumbnailing
Structure analysis

Which song?
Audio identification

Similar songs?
Cover song identification
Audio similarity search

Similar artists?
Genre classification
Music recommendation

Motivation

Music consumption 2020

I want to play that!
Automatic transcription
Melody extraction
Chord recognition
Key determination

Which content?
Speech / music / noise?
Audio classification

Spotify DEEZER
amazon music MUSIC

What is representative?
Audio Thumbnailing
Structure analysis

Which song?
Audio identification

Similar songs?
Cover song identification
Audio similarity search

Similar artists?
Genre classification
Music recommendation

Motivation

Music consumption 2020

40 000 000 Tracks

53 000 000 Tracks

30 000 000 Tracks

50 000 000 Tracks

Similar songs?
Cover song identification
Audio similarity search

Similar artists?
Genre classification
Music recommendation

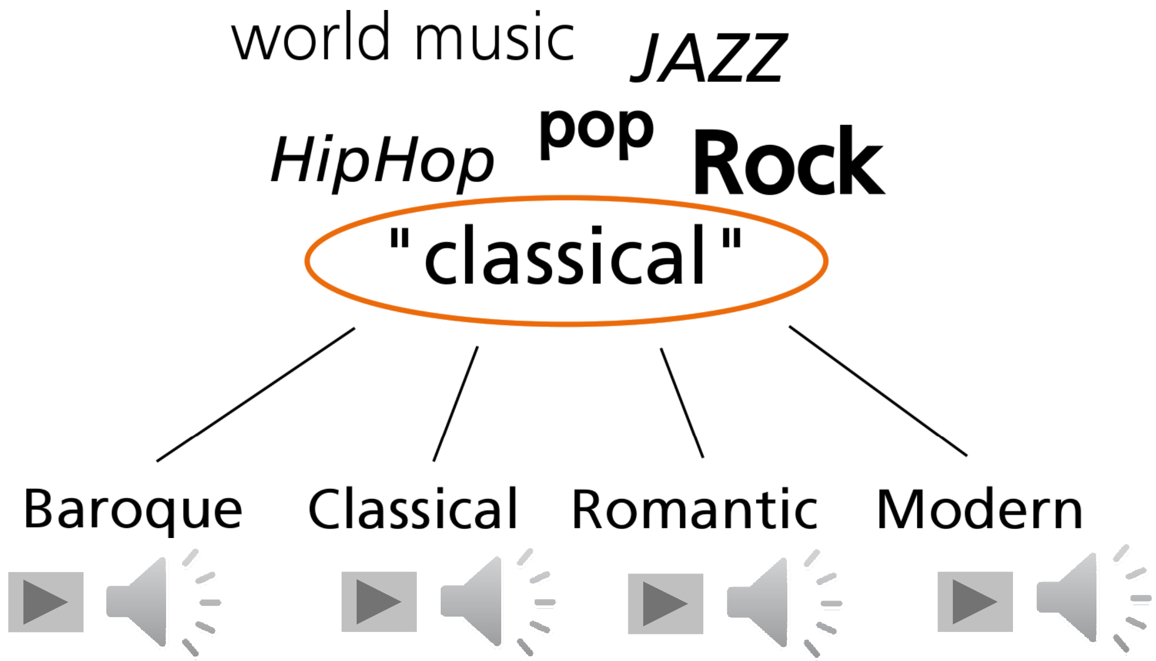
Music Genre Classification

world music JAZZ

HipHop pop **Rock**

"classical"

Music Genre Classification



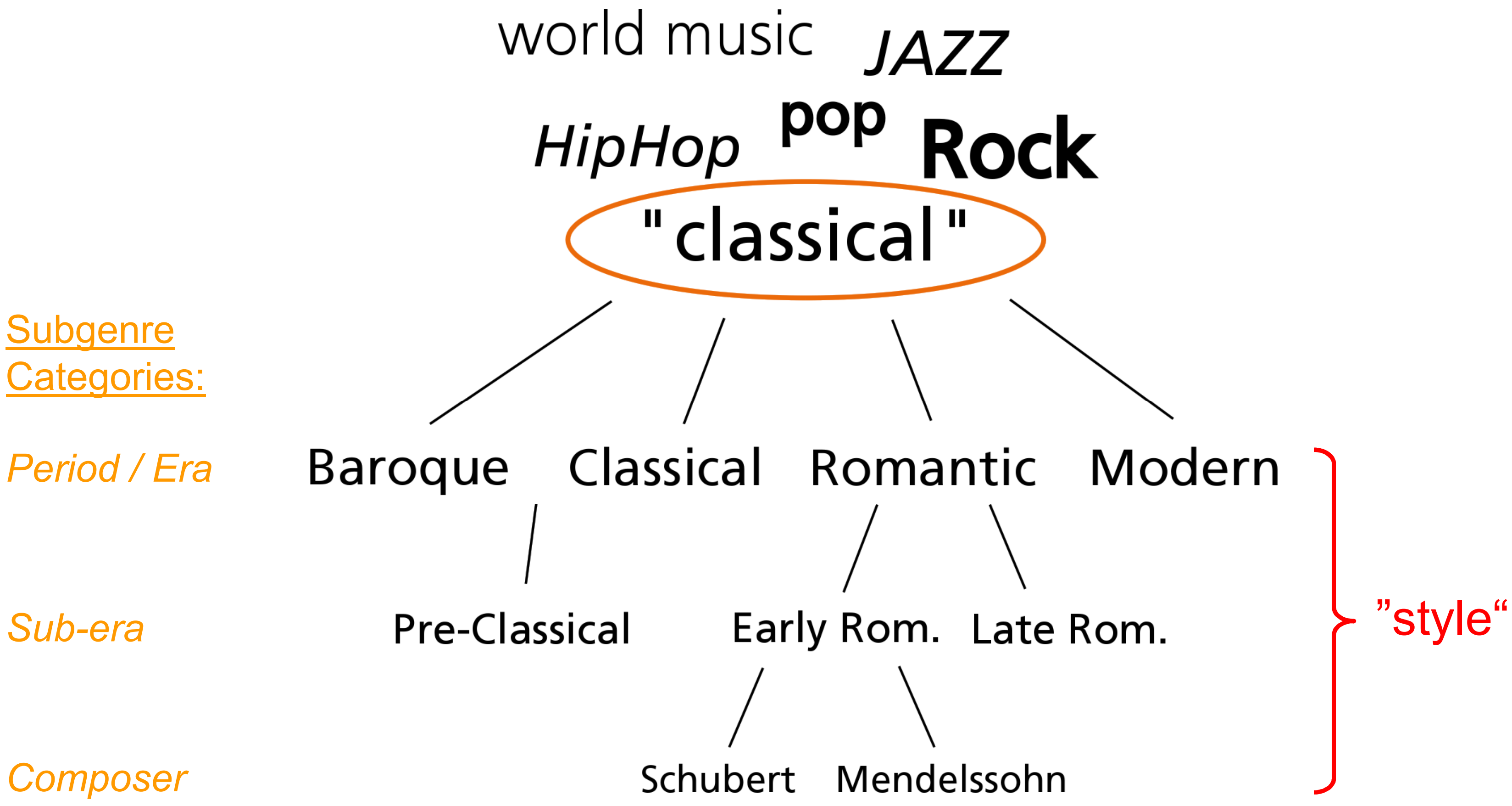
J. S. Bach,
Brandenburg Concerto
No. 2 in F major, I. Allegro,
Cologne Chamber Orch.

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

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

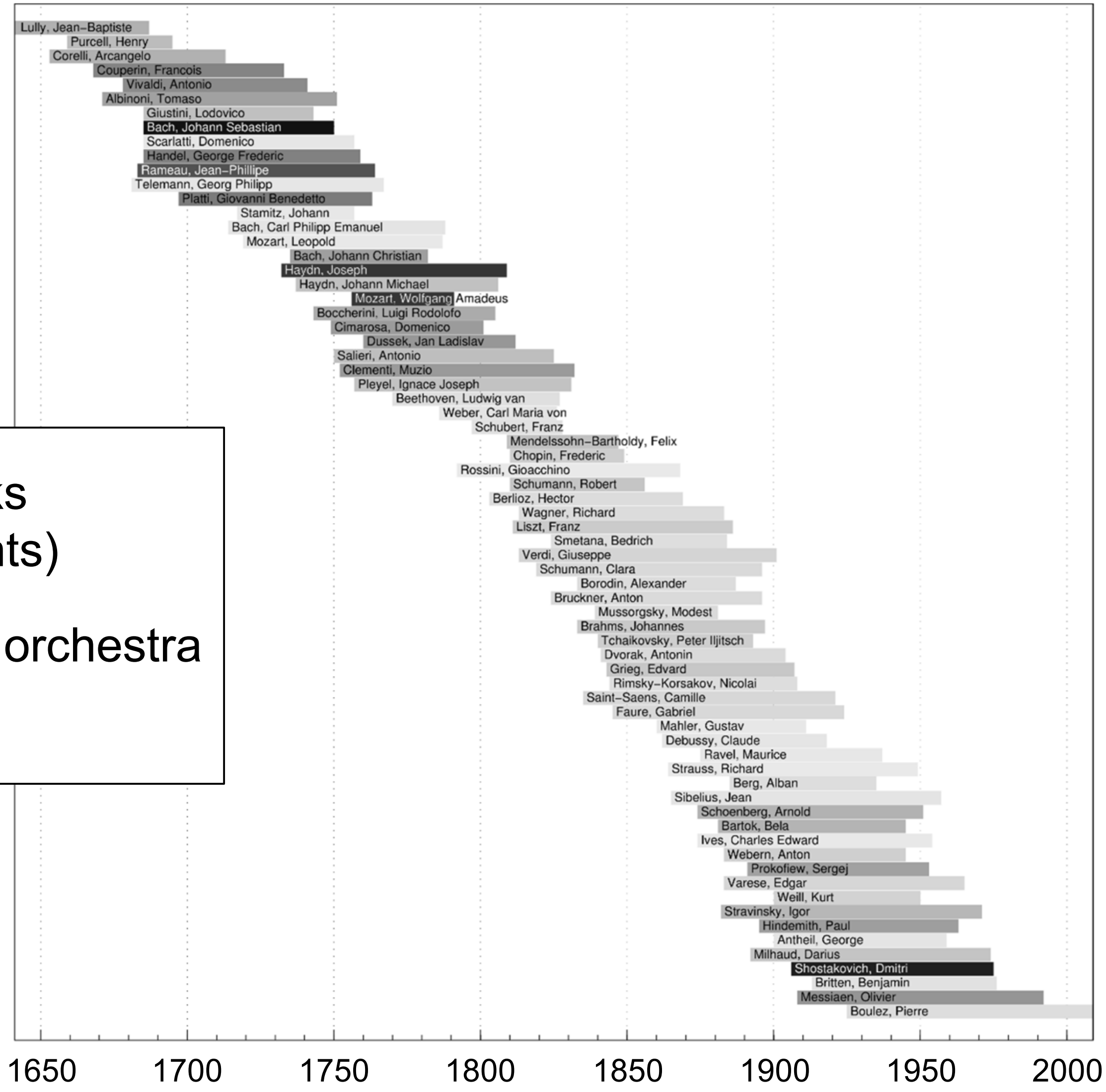
A. Webern,
Variations for Orchestra op. 30
Ulster Orchestra

Music Genre Classification

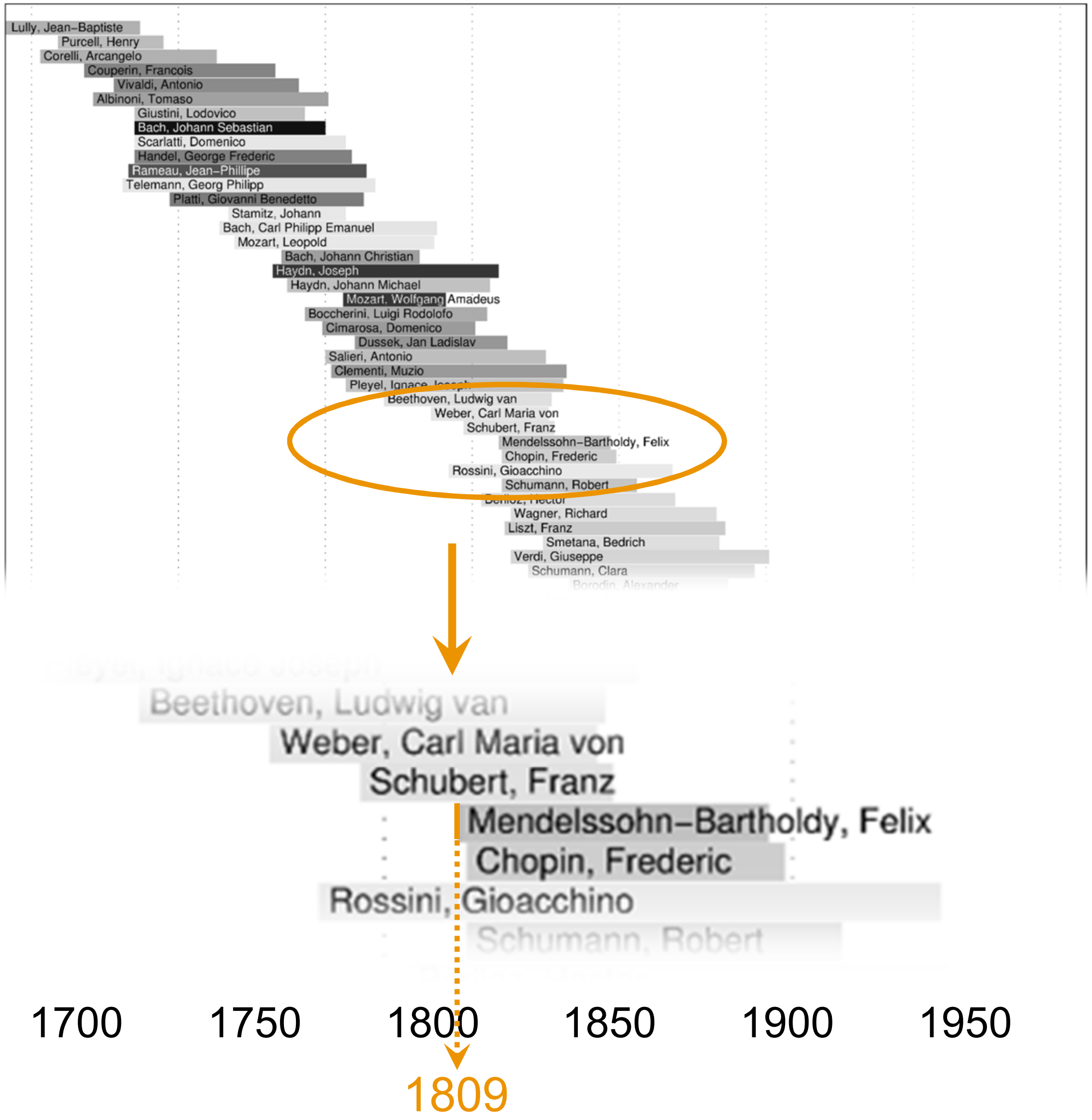


Style Classification: Dataset

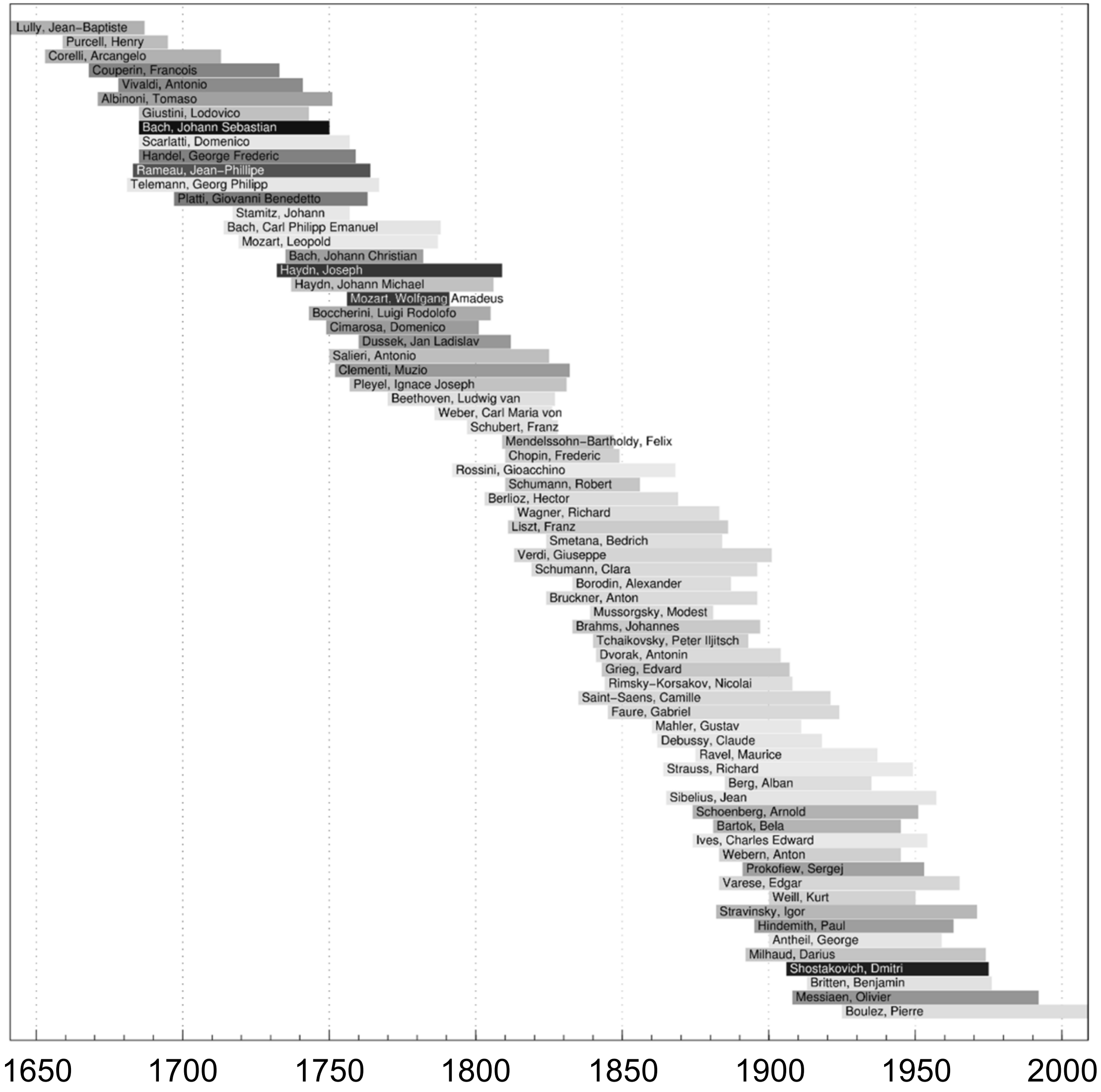
- 2000 tracks (movements)
- piano and orchestra balanced



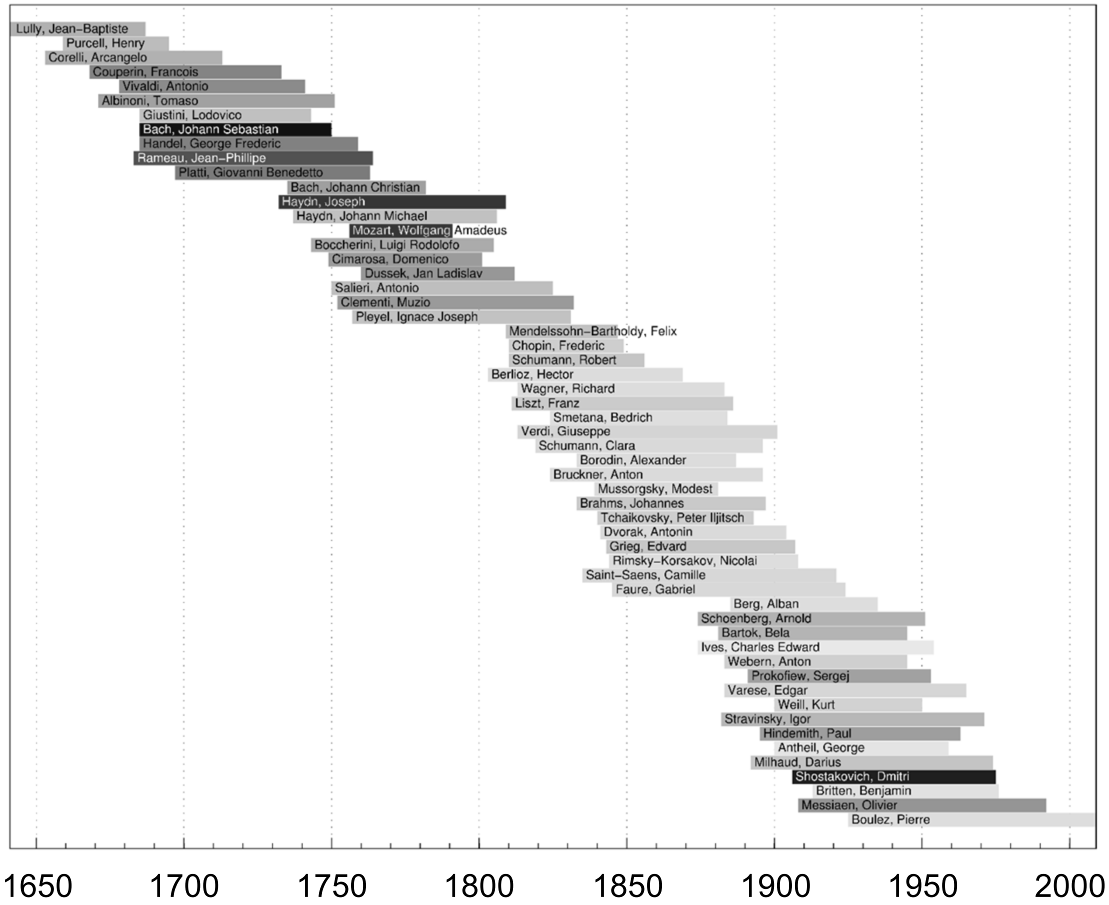
Style Classification: Dataset



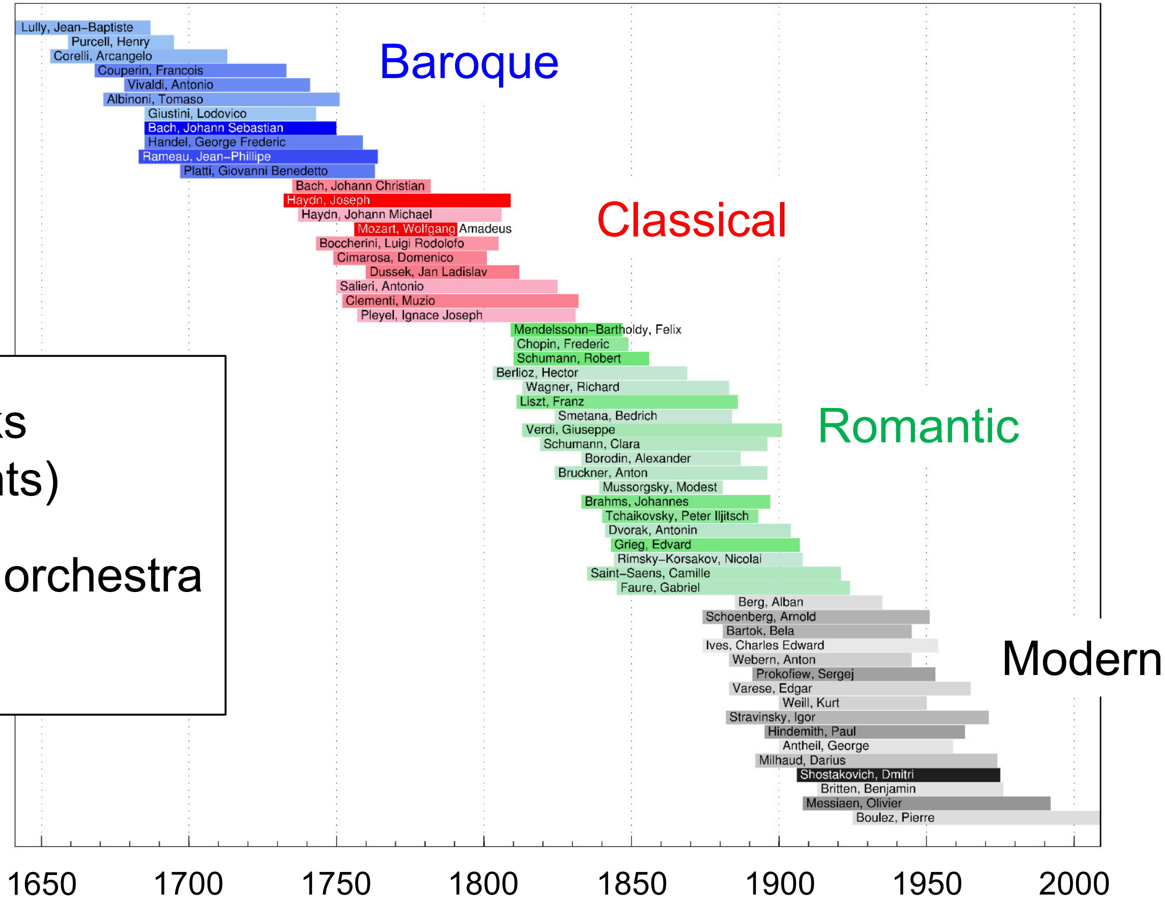
Style Classification: Dataset



Style Classification: Eras



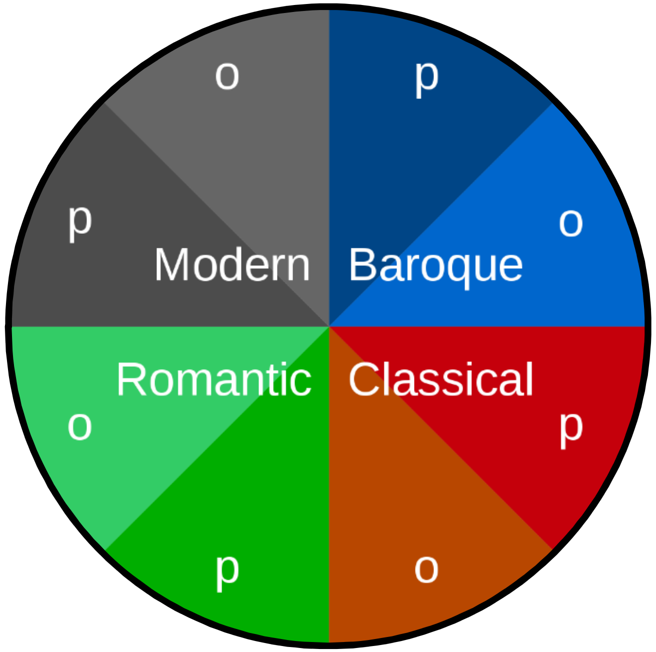
Style Classification: Eras



- 1600 tracks (movements)
- piano and orchestra balanced

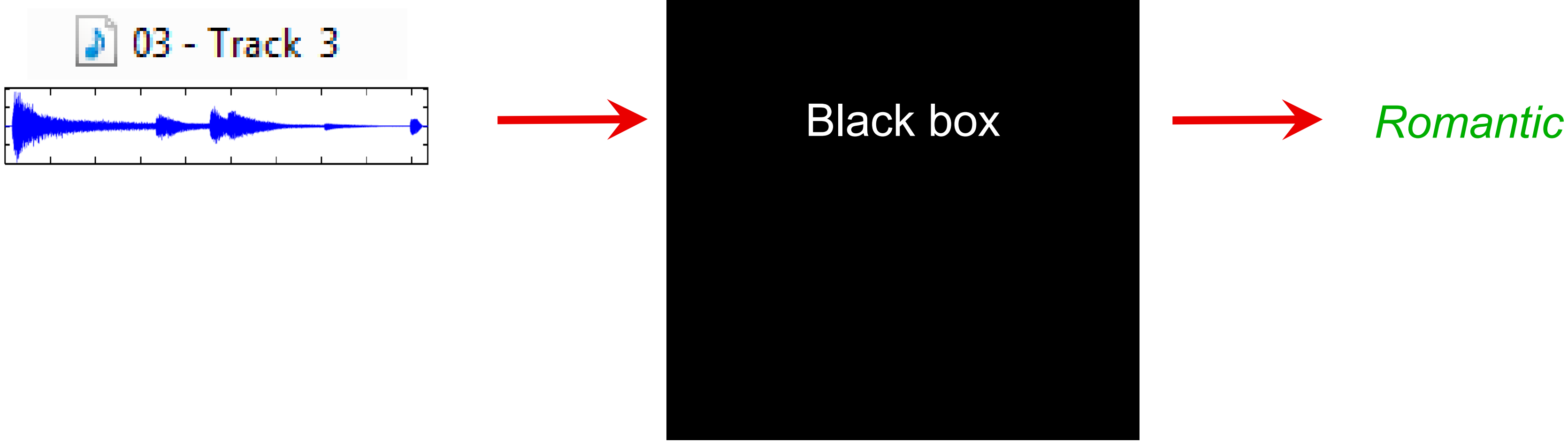
Style Classification: Eras

- Balanced: 800 piano tracks (p), 800 orchestra tracks (o)
- Each 200 tracks → 1600 in total

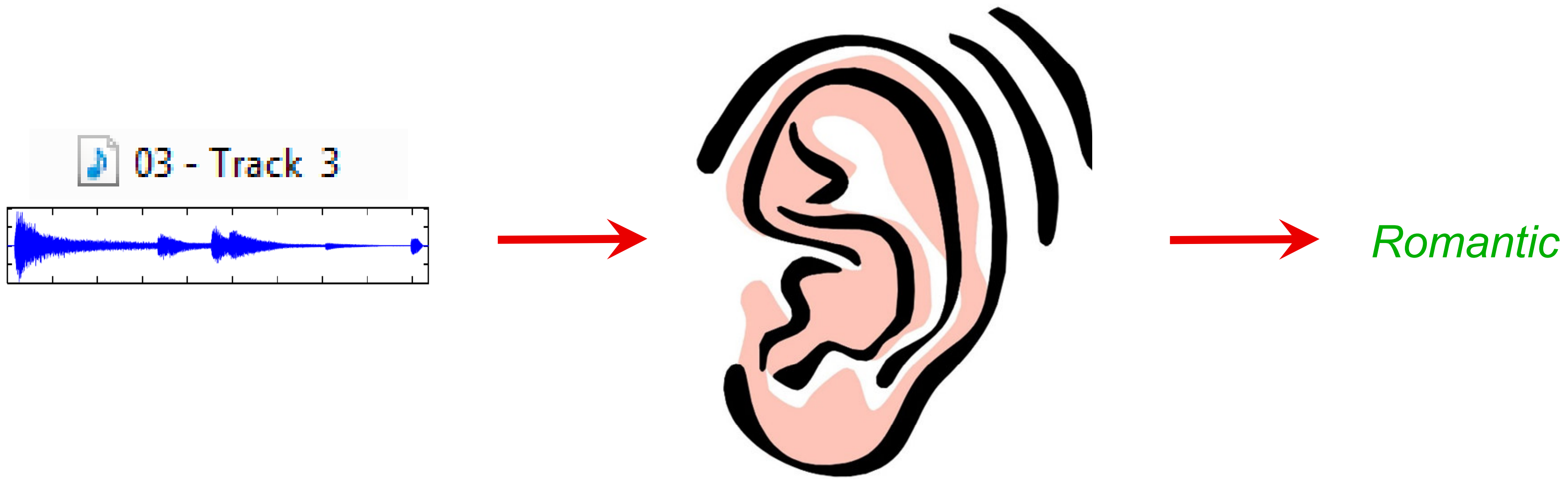


Classification problem
4-class problem

Style Classification: Machine Learning



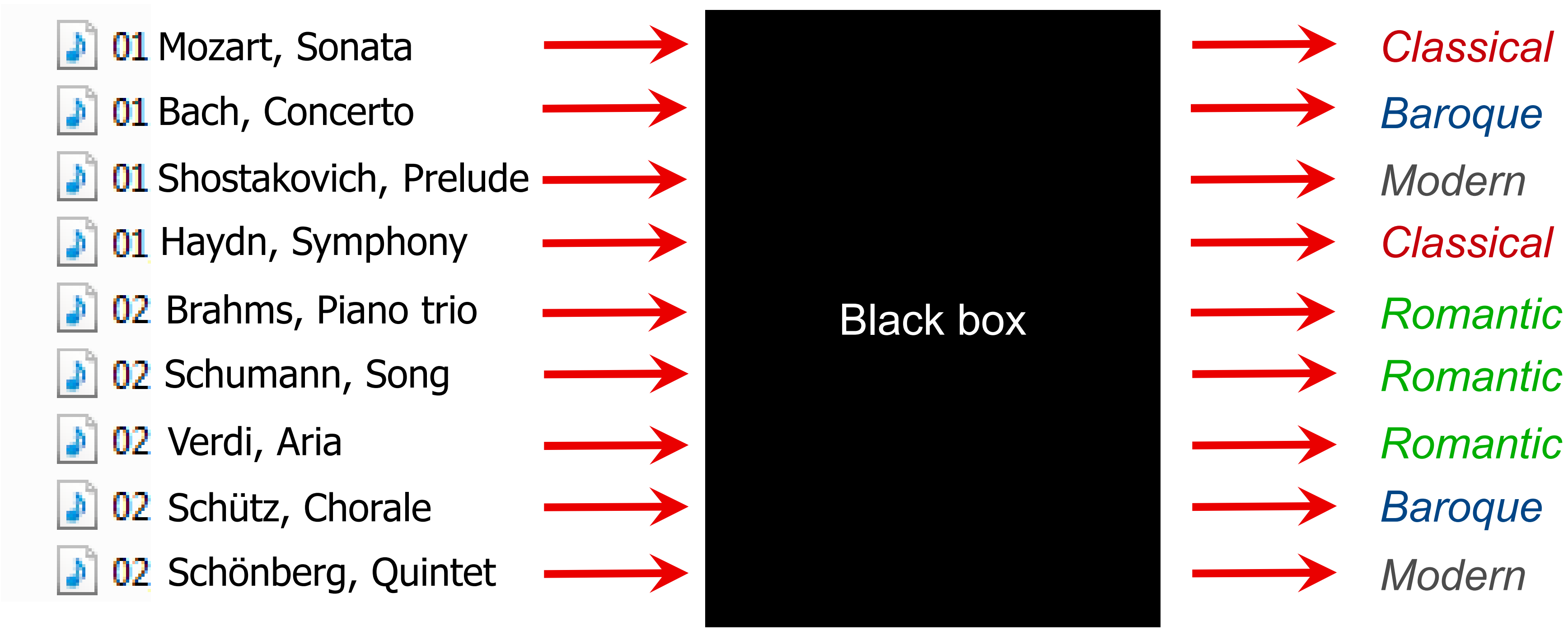
Style Classification: Machine Learning



Style Classification: Machine Learning

01 Mozart, Sonata	*Classical*
01 Bach, Concerto	*Baroque*
01 Shostakovich, Prelude	*Modern*
01 Haydn, Symphony	*Classical*
02 Brahms, Piano trio	*Romantic*
02 Schumann, Song	*Romantic*
02 Verdi, Aria	*Romantic*
02 Schütz, Chorale	*Baroque*
02 Schönberg, Quintet	*Modern*

Style Classification: Machine Learning

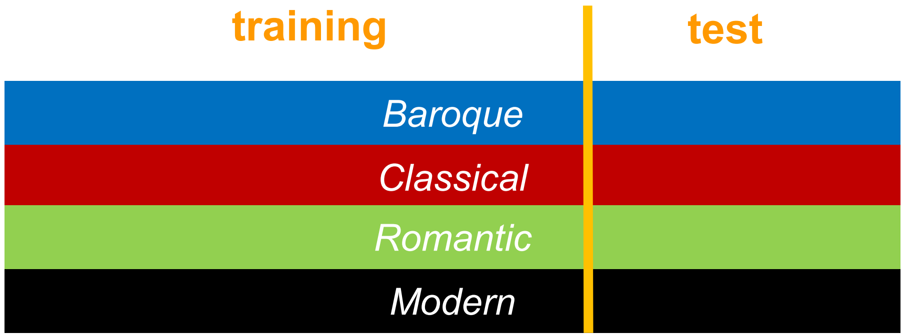


Style Classification: Machine Learning

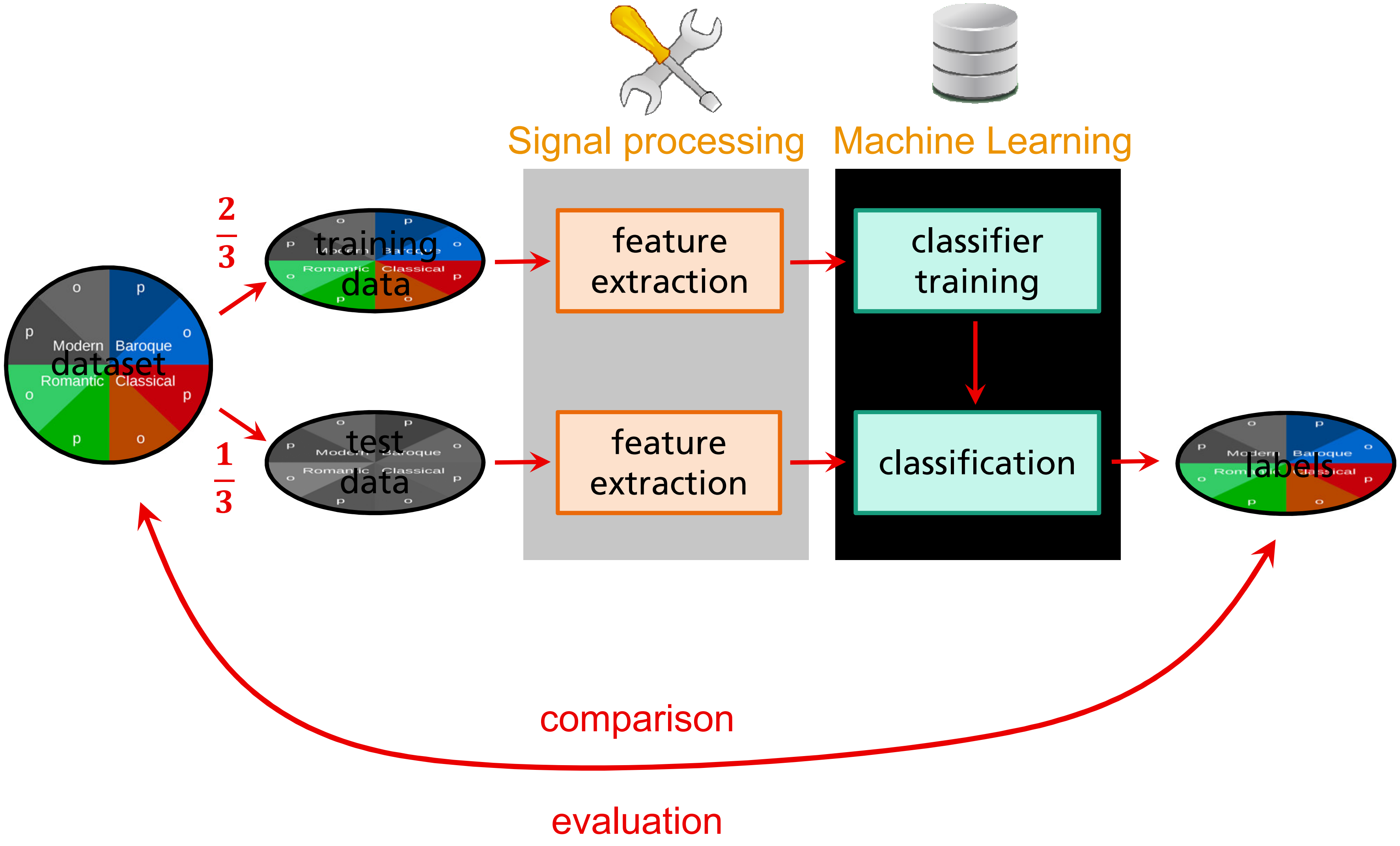
- Experimental design: Evaluation with **Cross Validation** (CV)
- Separate data into different parts (***folds***)

	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



Style Classification: Machine Learning



Style Classification: Feature extraction

Signal processing

feature extraction

feature extraction

Style Classification: Feature extraction

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

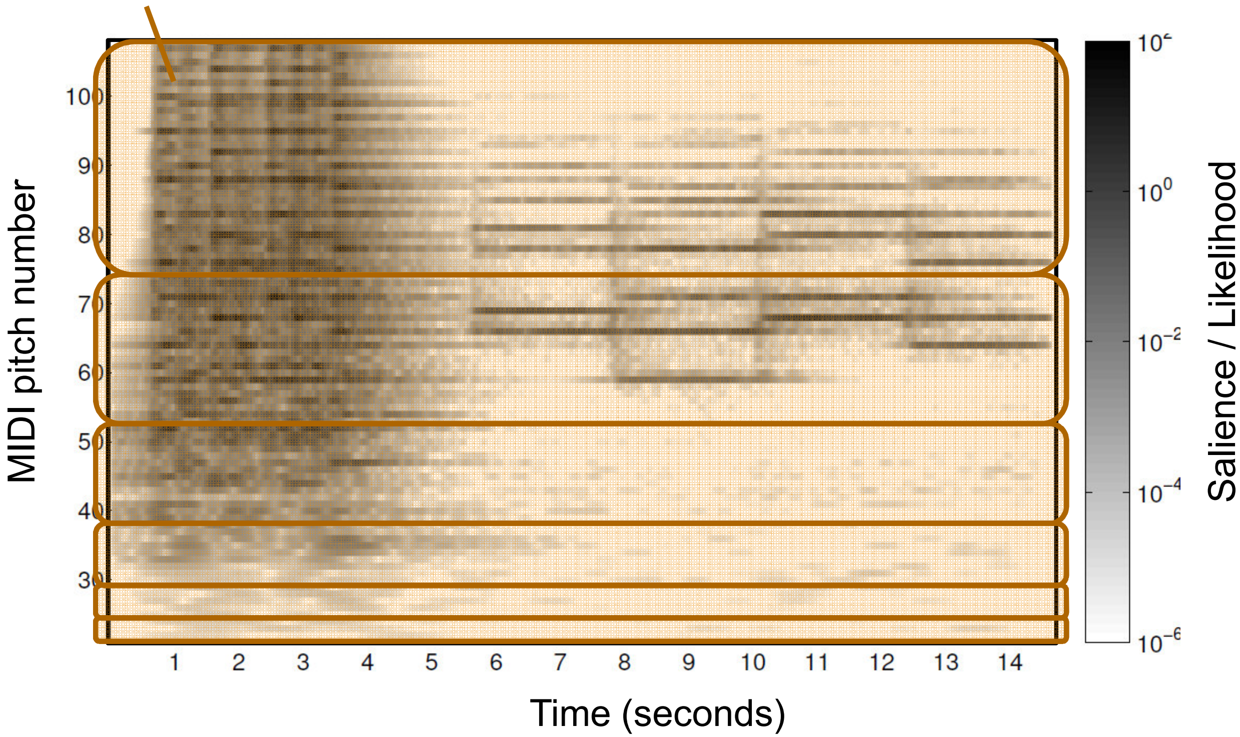
Recall: Spectral Features

→ independent of exact pitches

→ describe **timbral** properties (sound color)

„standard features“ for genre classification

Frequency bands: Loudness, Spectral Flatness, Spectral Centroid



Style Classification: Feature extraction

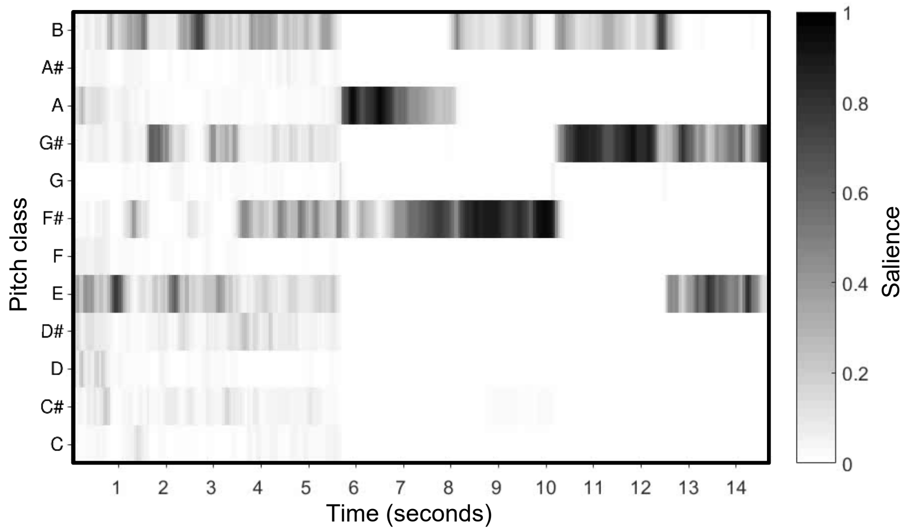
- Standard approach (*content-based*)
  - Supervised machine learning
  - Based on spectral / timbral features

- In classical music → Instrumentation

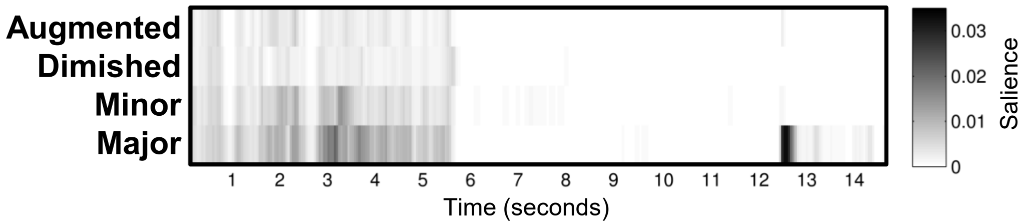
- Better categories?
  - *Musical style*
  - Independent from instrumentation
  - → **Tonality / Harmony**

Recall: Chord Type and Interval Features

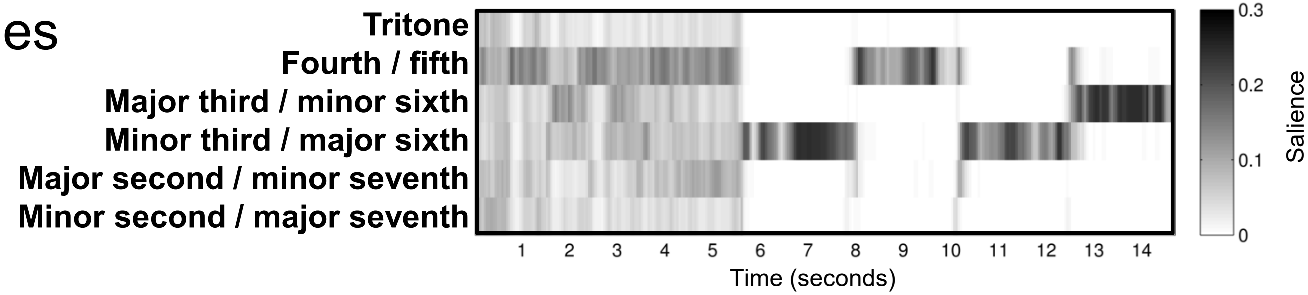
- Chromagram



- Chord types



- Interval categories

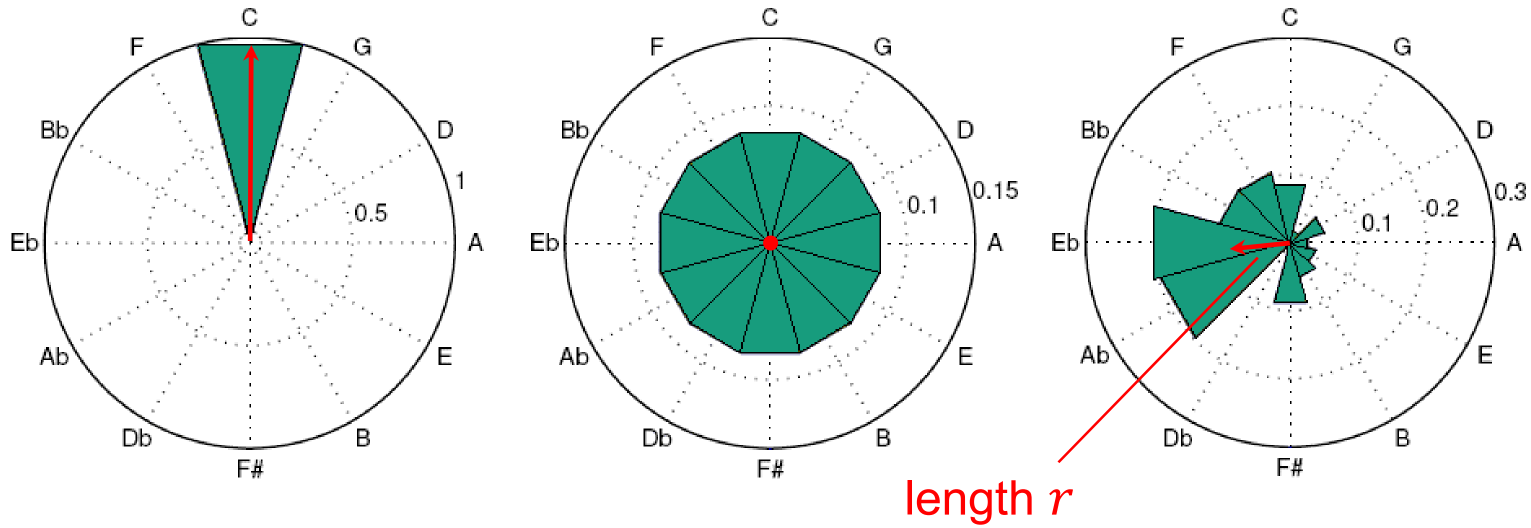


→ transposition-invariant features!

Recall: Tonal Complexity Features

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

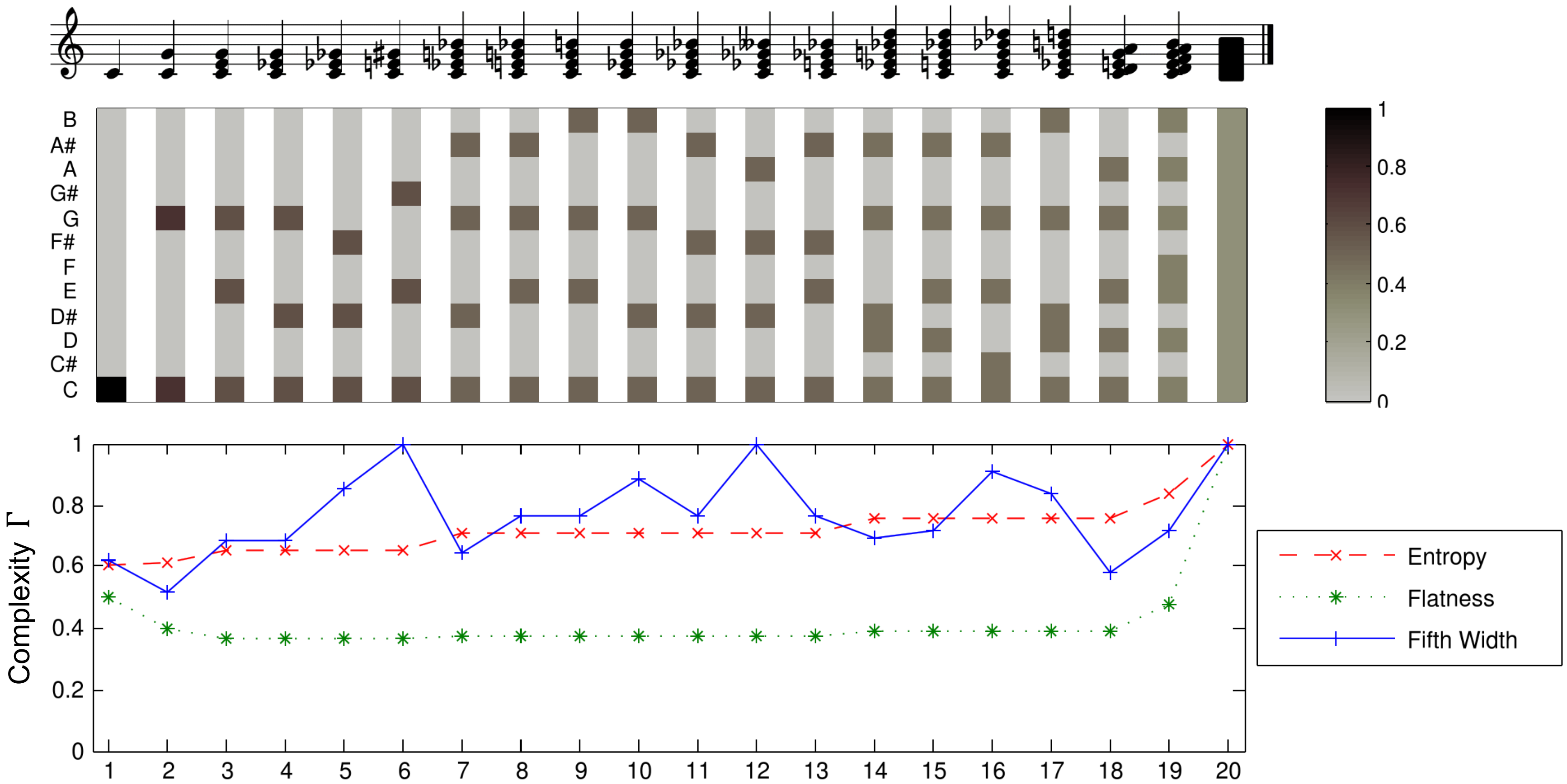
$\Gamma = 0$ $\qquad$ $\Gamma = 1$ $\qquad$ $0 < \Gamma < 1$



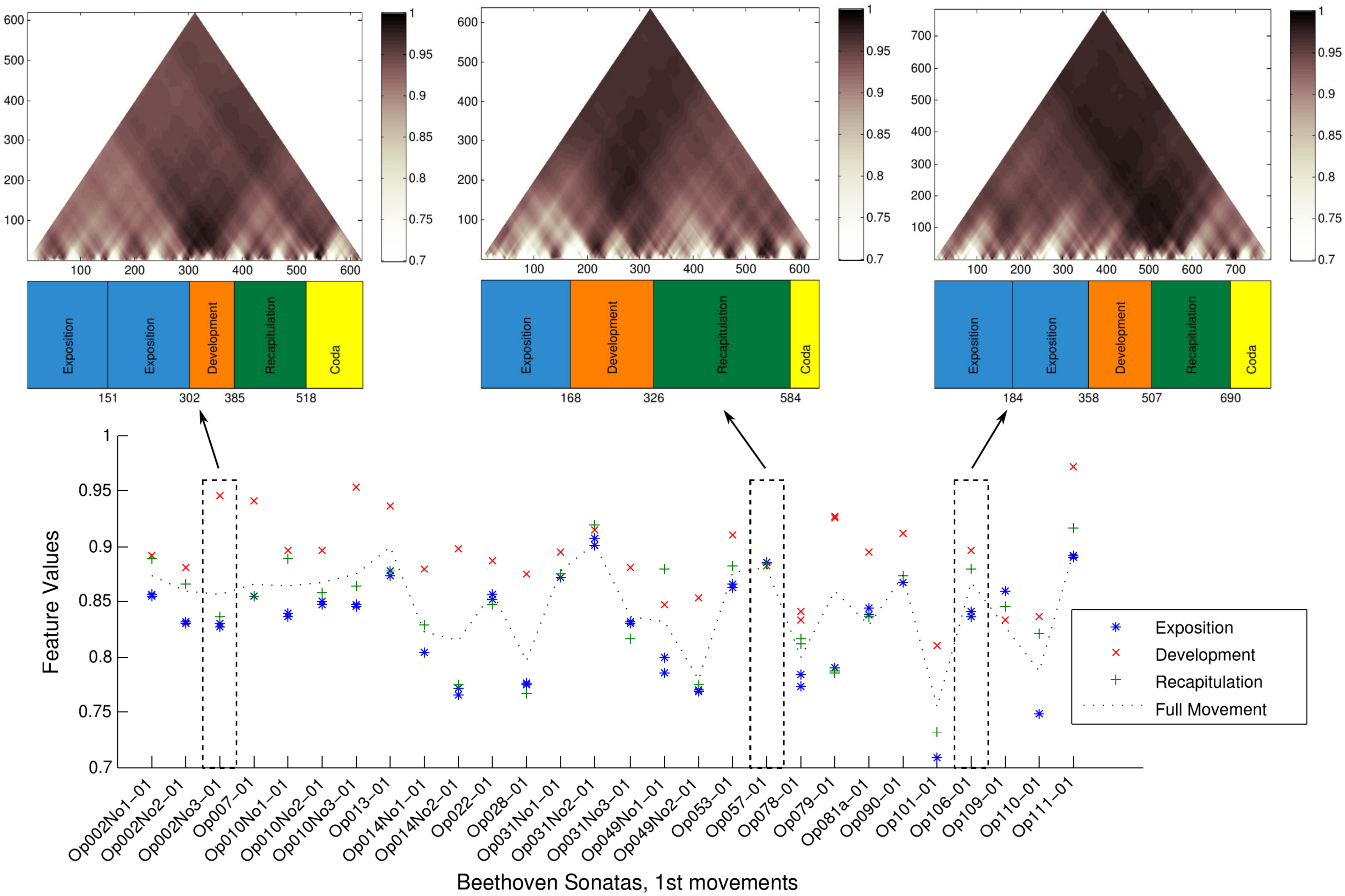
$$\boxed{\Gamma = \sqrt{1 - r}}$$

- Relating to different time scales!

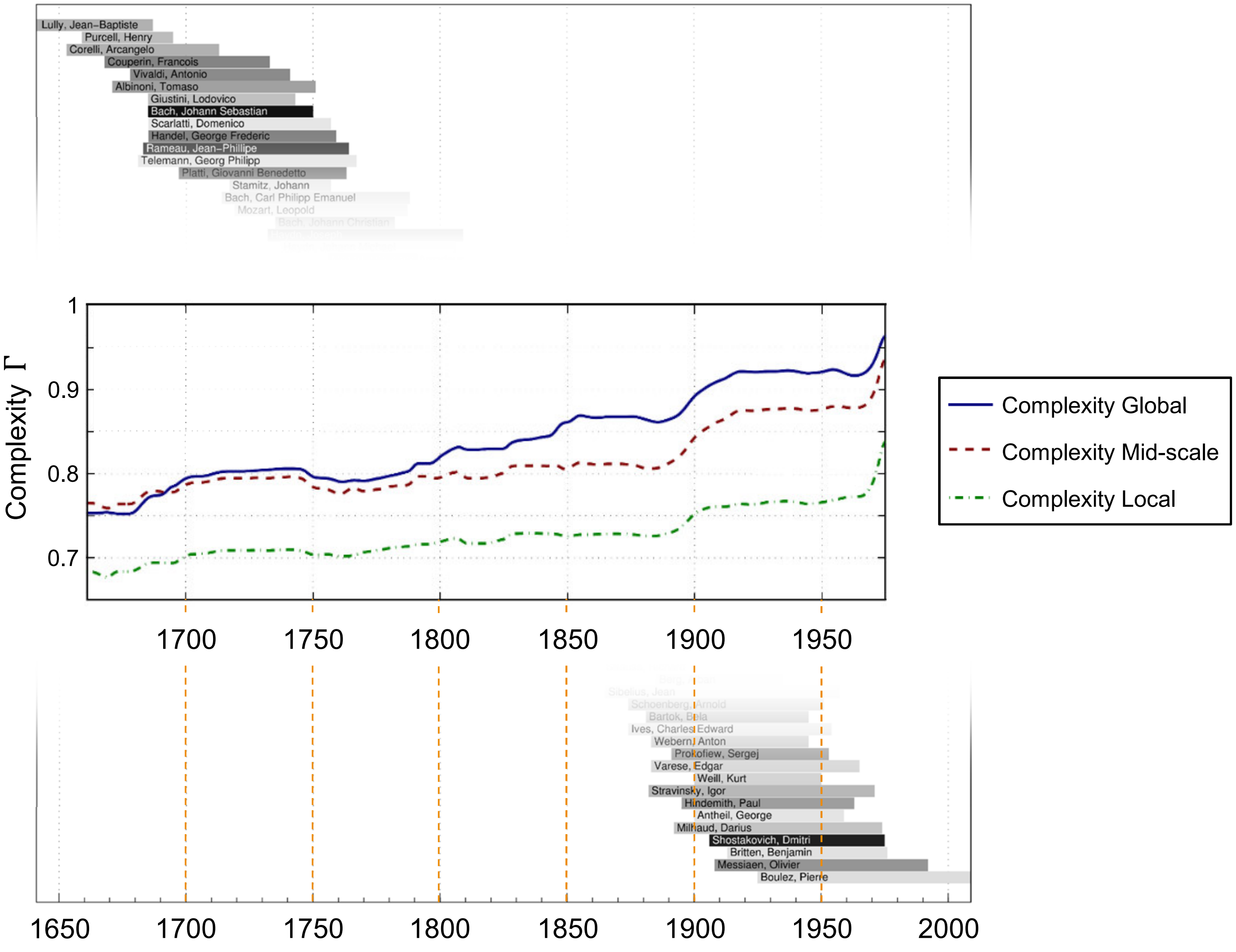
Recall: Tonal Complexity Features



Recall: Tonal Complexity Features



Recall: Tonal Complexity



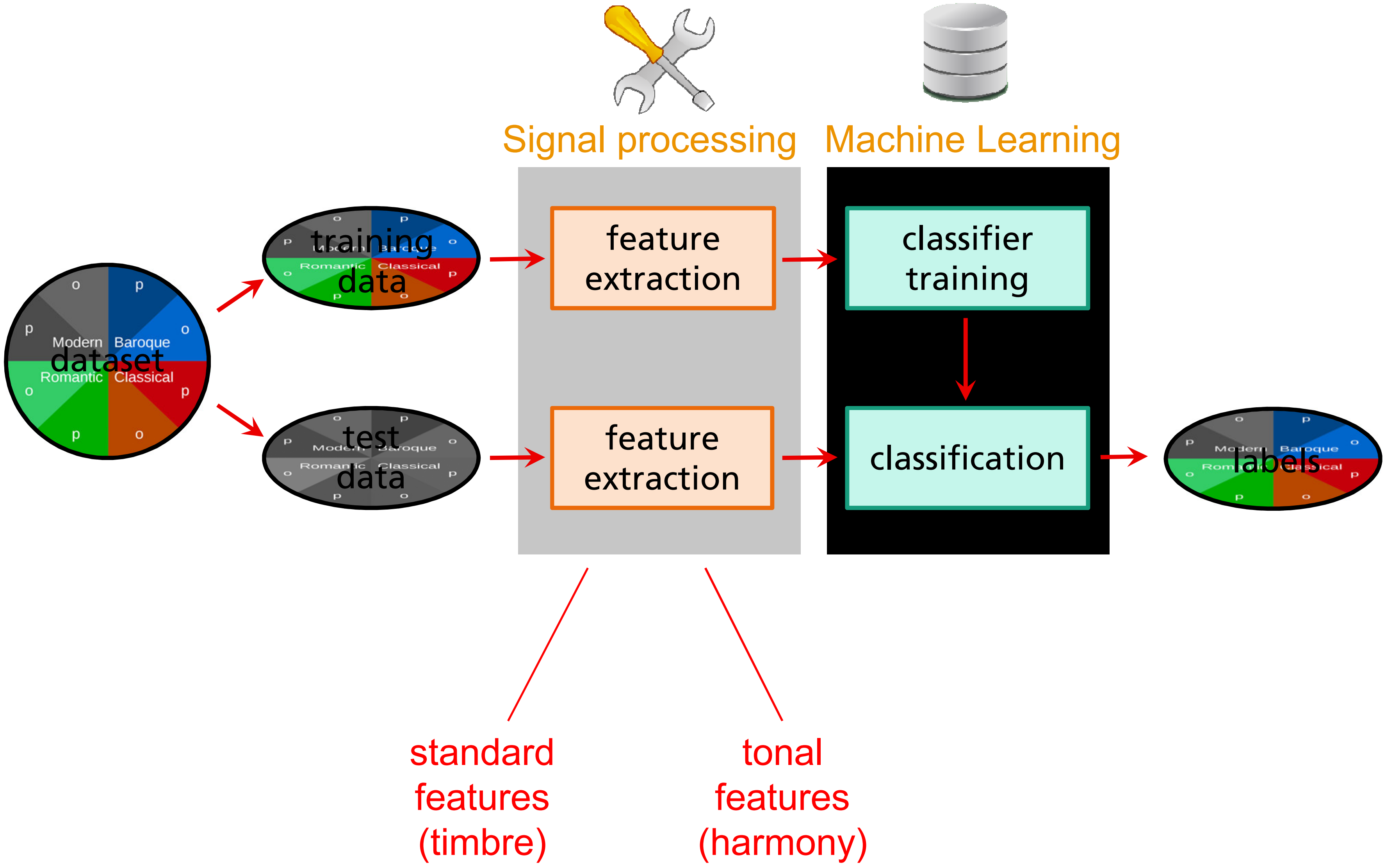
Style Classification: Feature extraction

12…16 frequency bands

4 time scales

Standard	Dim.	Tonal	Dim.
MFCC	16	Interval categories	6 x 4
OSC	14	Chord types	4 x 4
ZCR	1	Complexity	7 x 4
ASE	16	Chord transitions	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**

Style Classification: Machine Learning

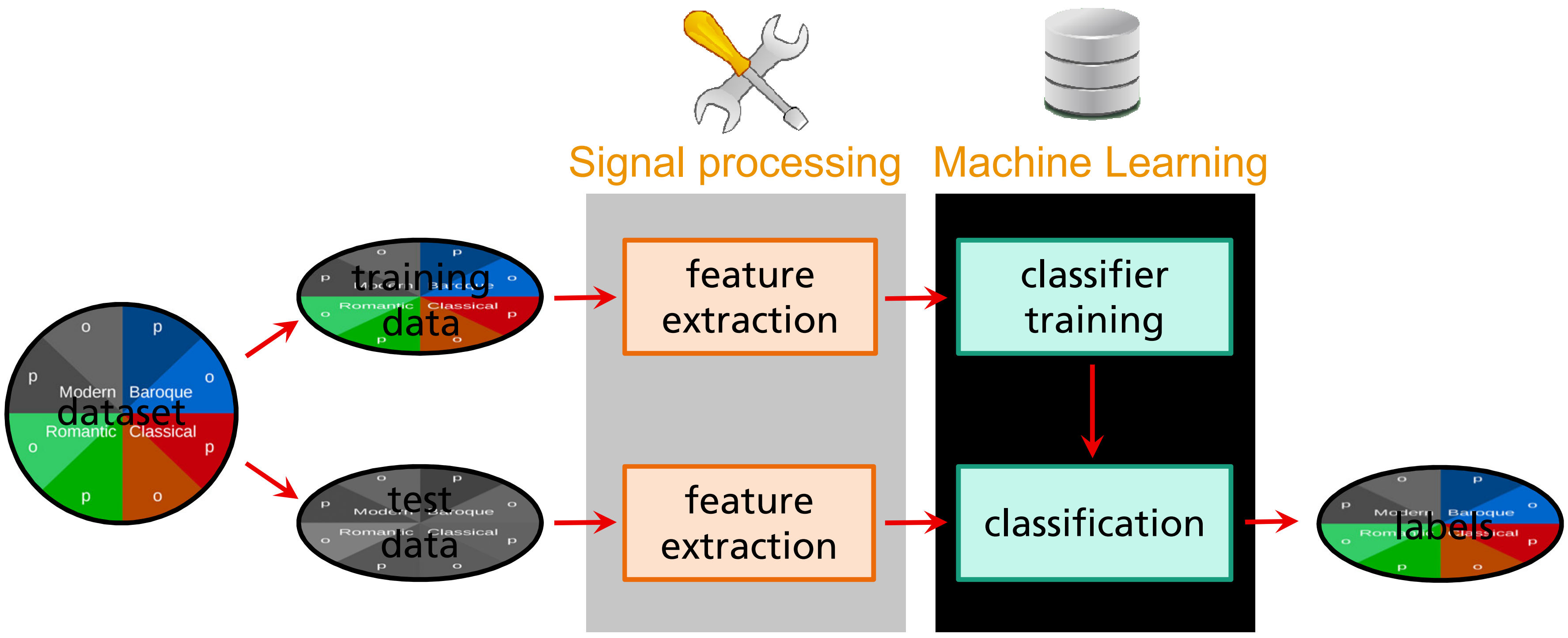


Style Classification
Overview

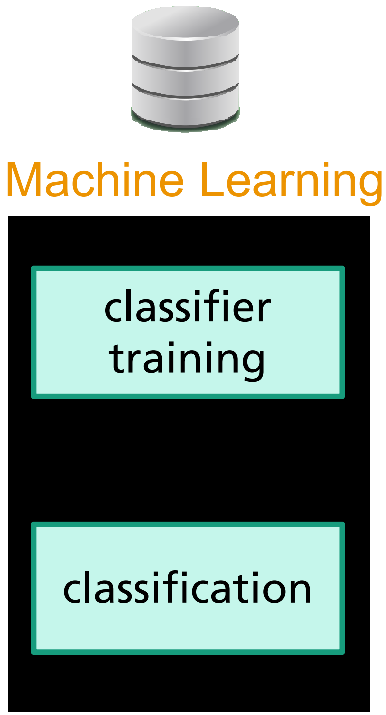
Machine Learning pipeline:

- Feature extraction
- Classification

Style Classification: Machine Learning

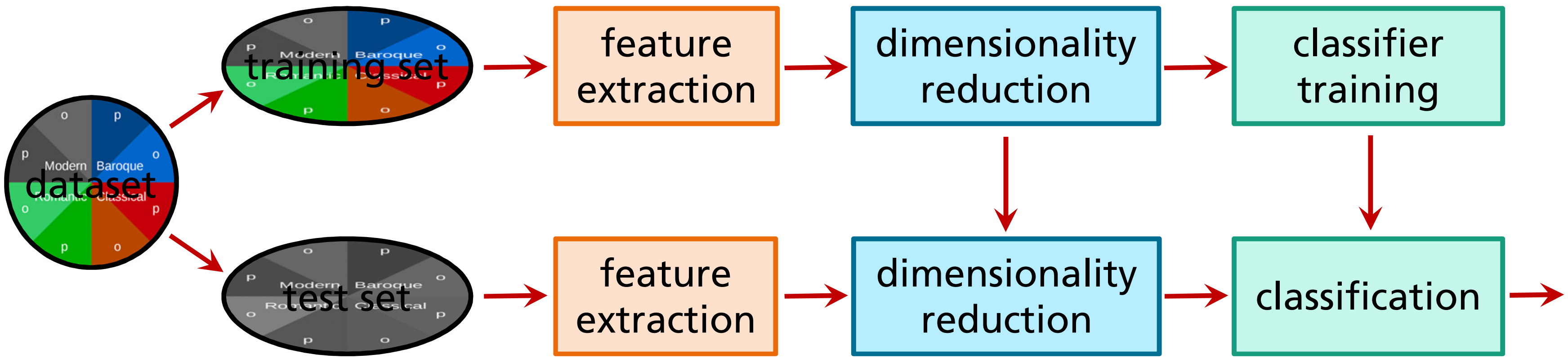


Style Classification: Machine Learning



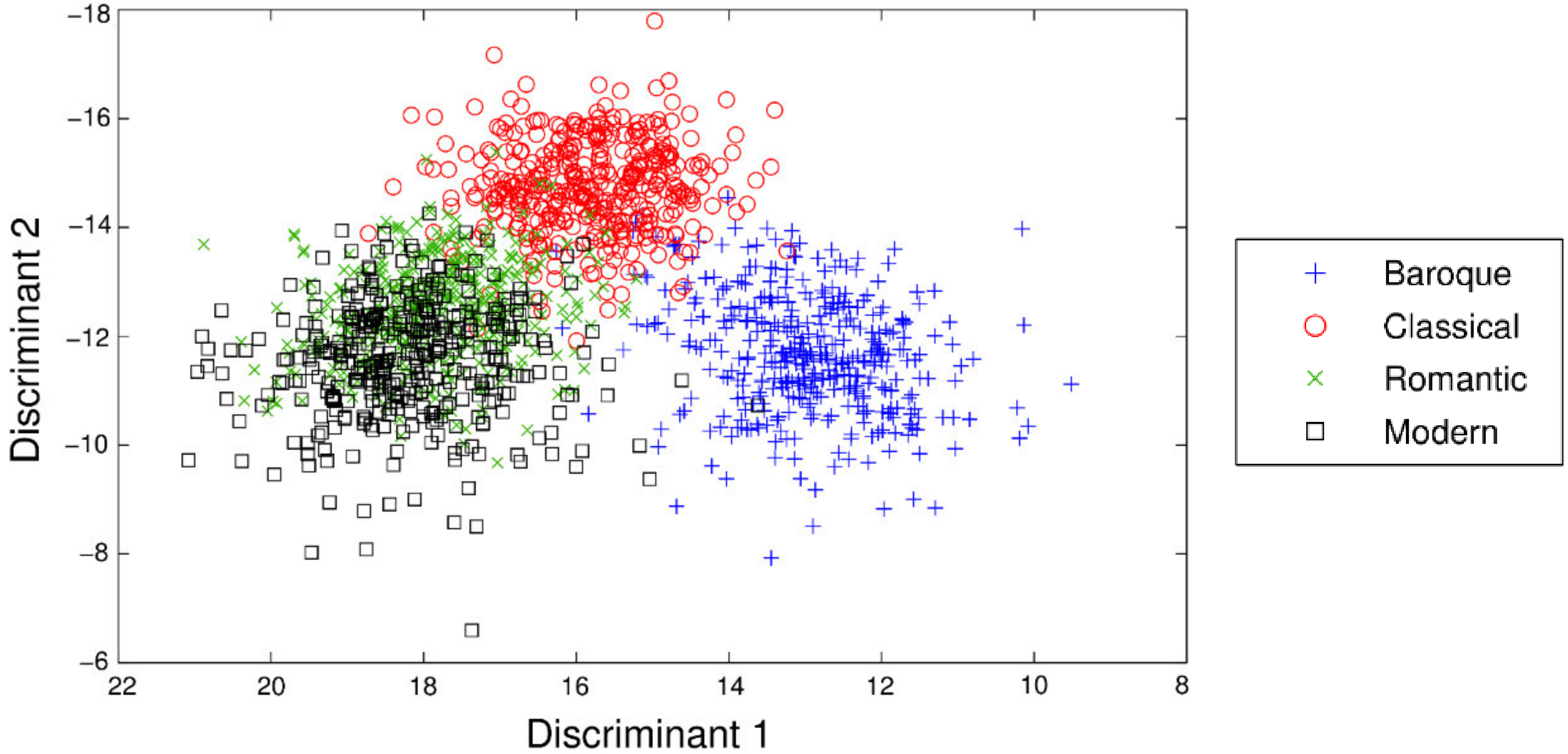
Style Classification: Machine Learning

- Supervised machine learning



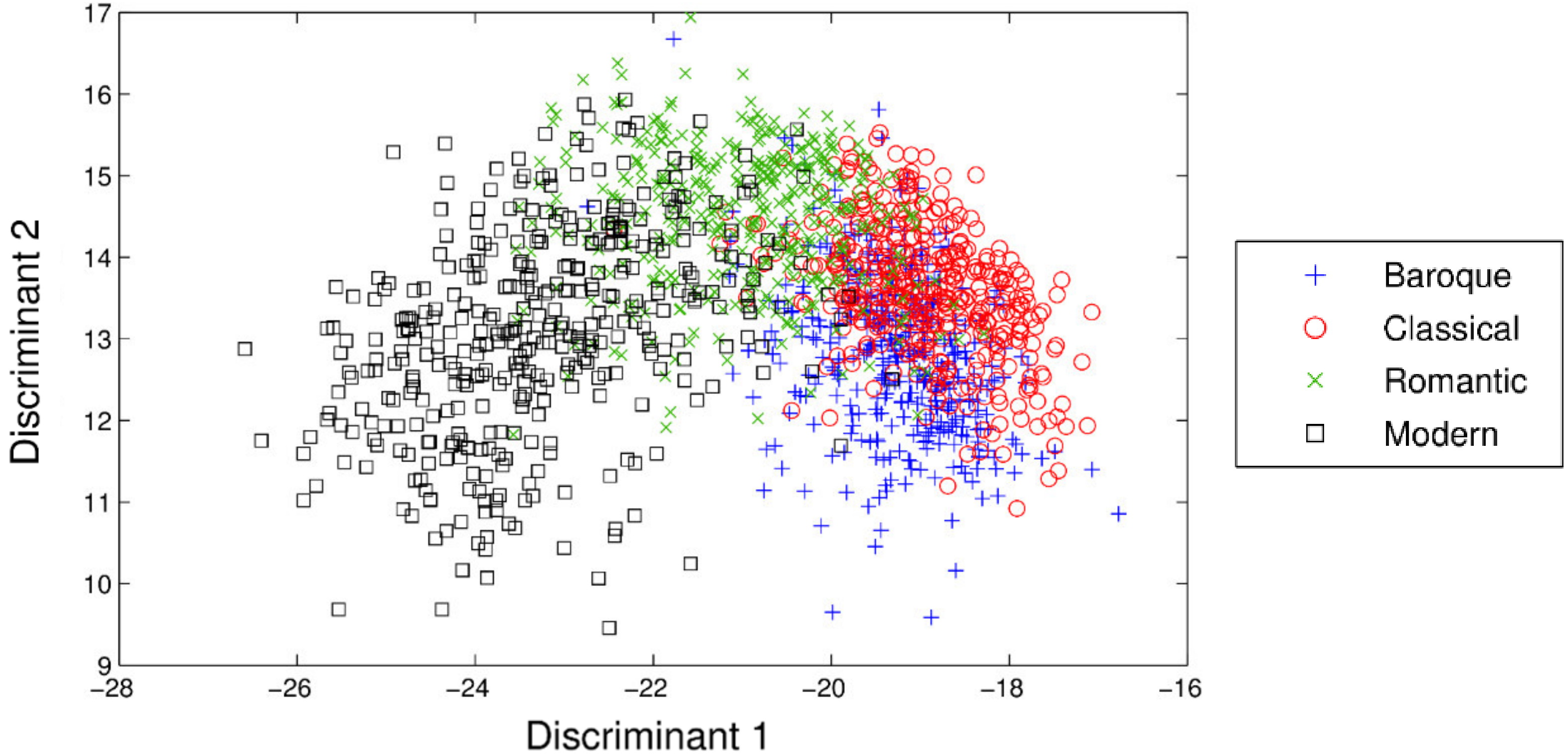
Dimensionality Reduction

- 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, …)



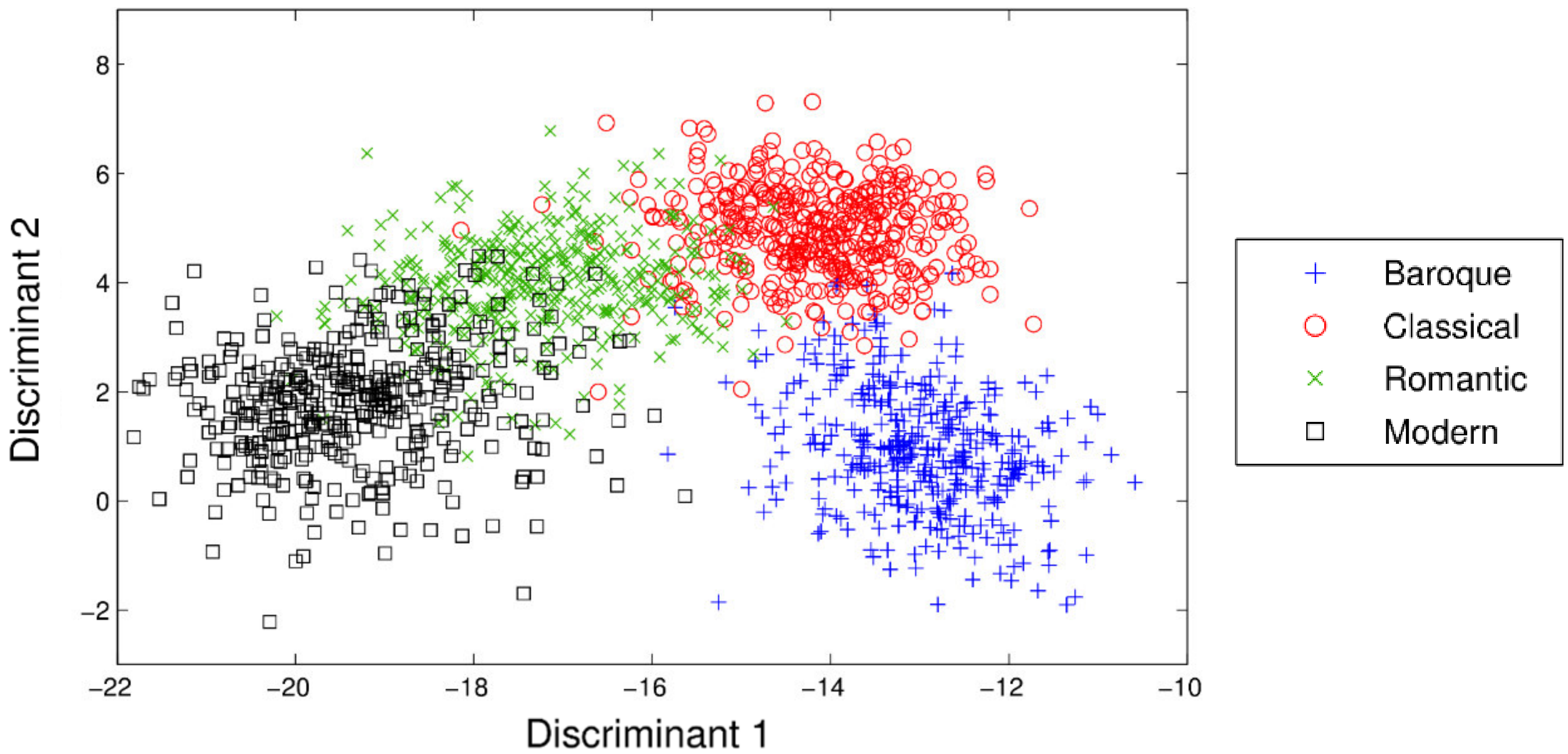
Dimensionality Reduction

- 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

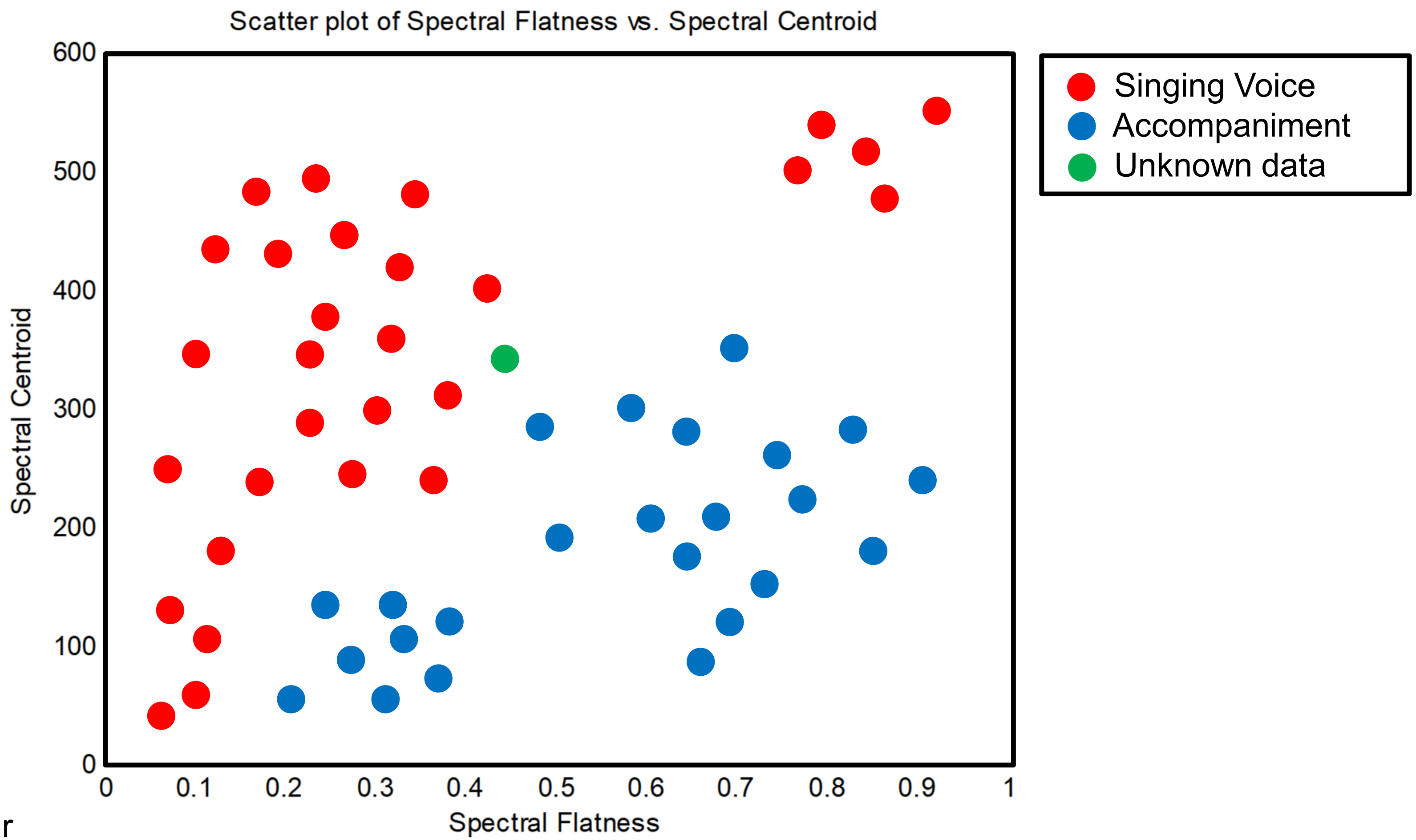
- Reduce feature space to few dimensions

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

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

Classification Methods

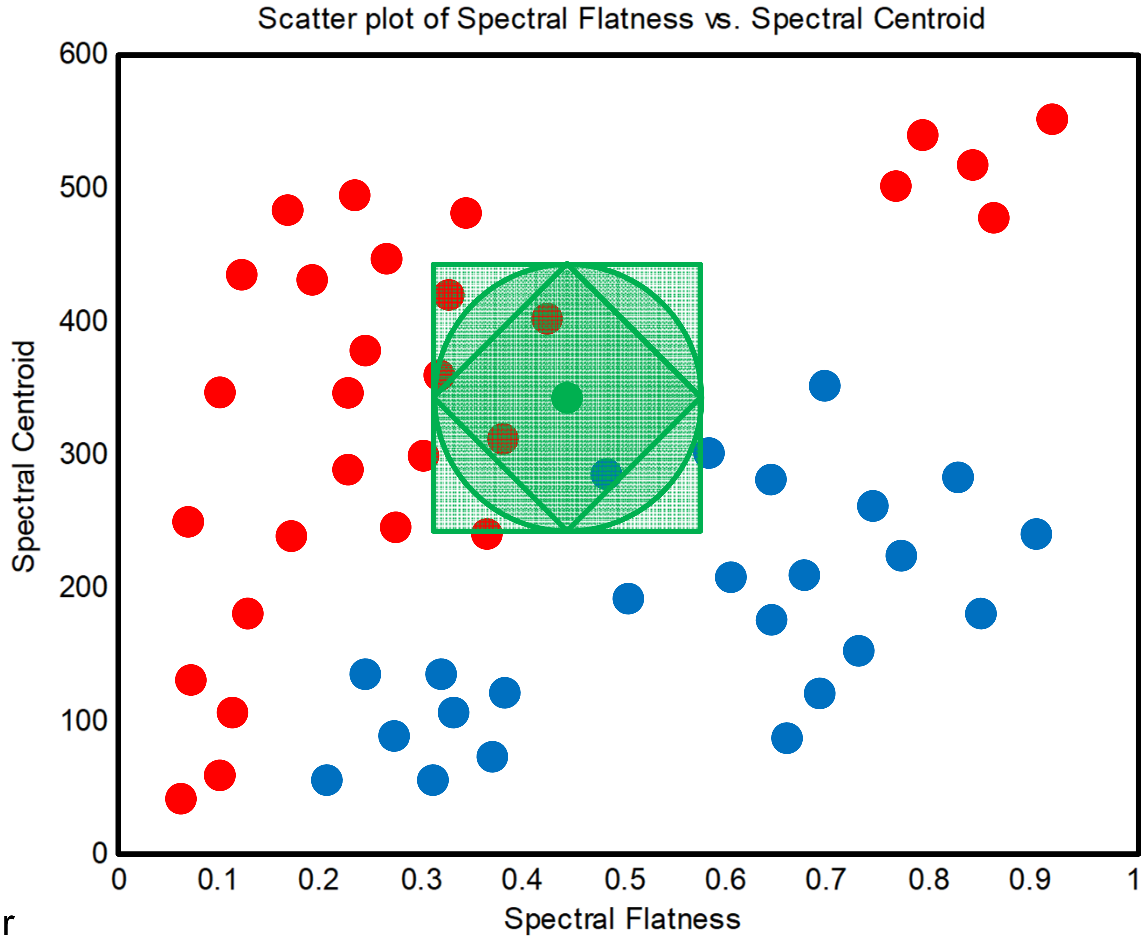
- *k* Nearest Neighbours (kNN)



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)

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

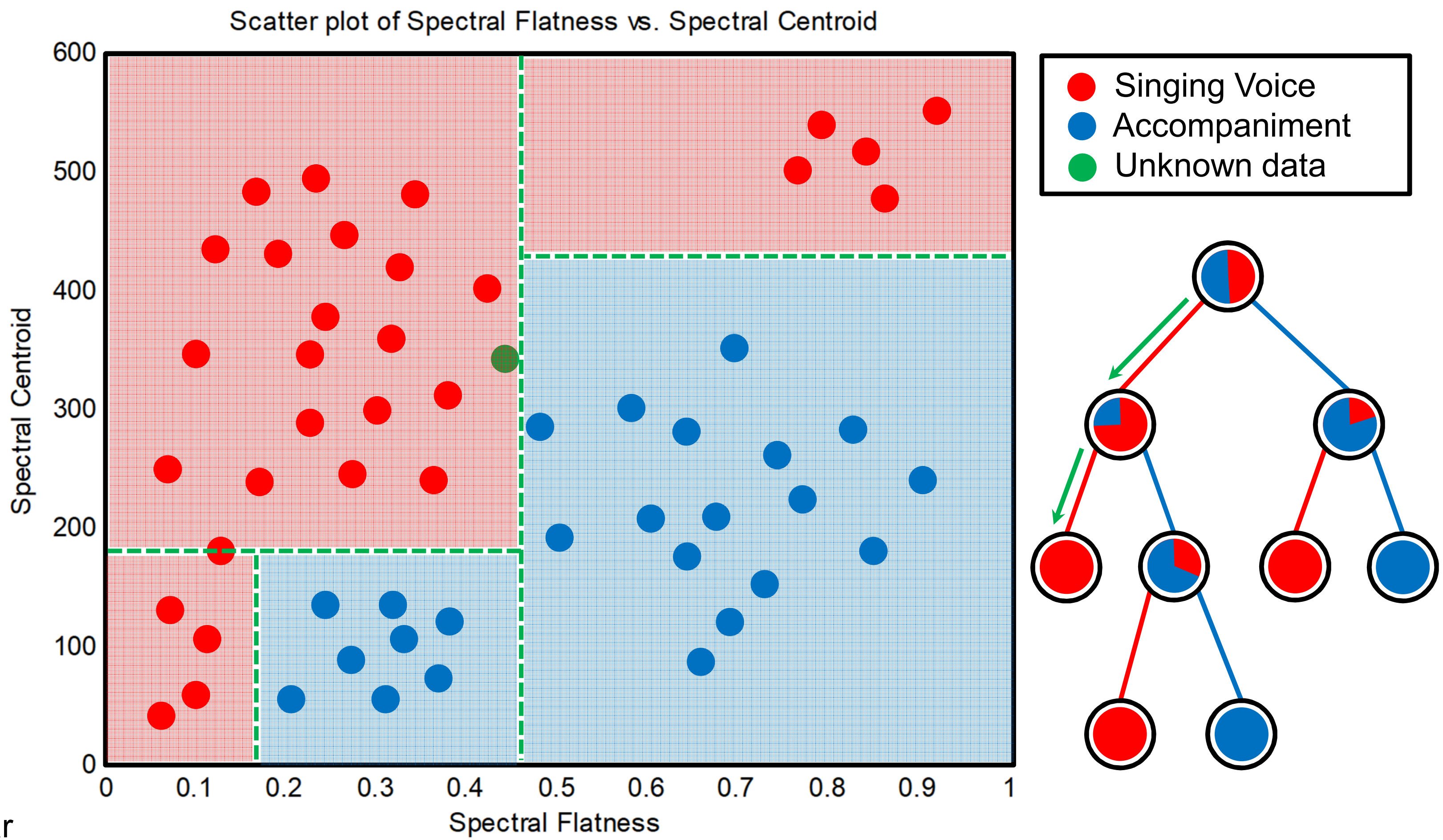
L∞-Dist. (Maximum)

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

Slides:
Christian Dittmar

Classification Methods

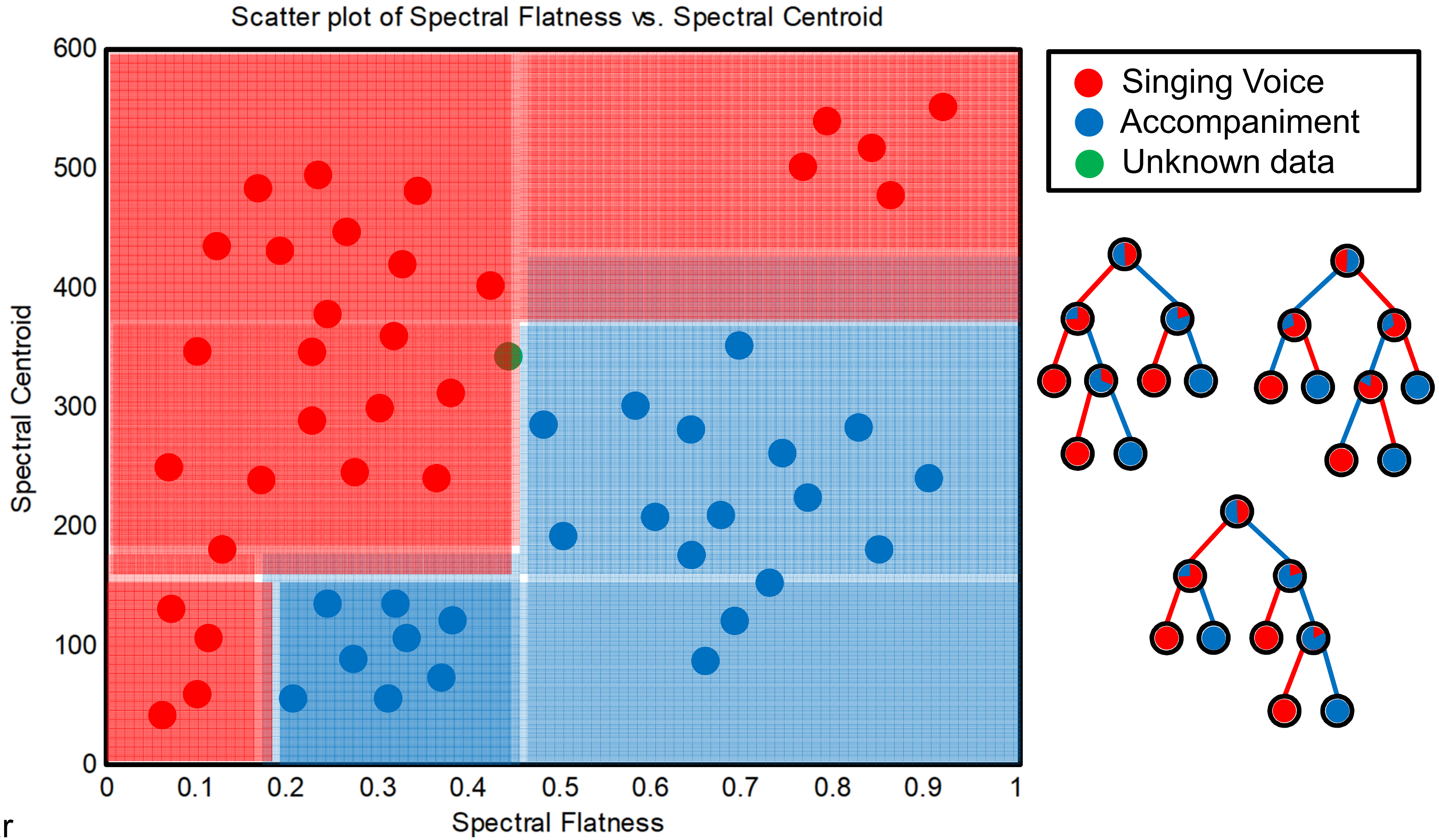
- Decision Trees (DT)



Slides:
Christian Dittmar

Classification Methods

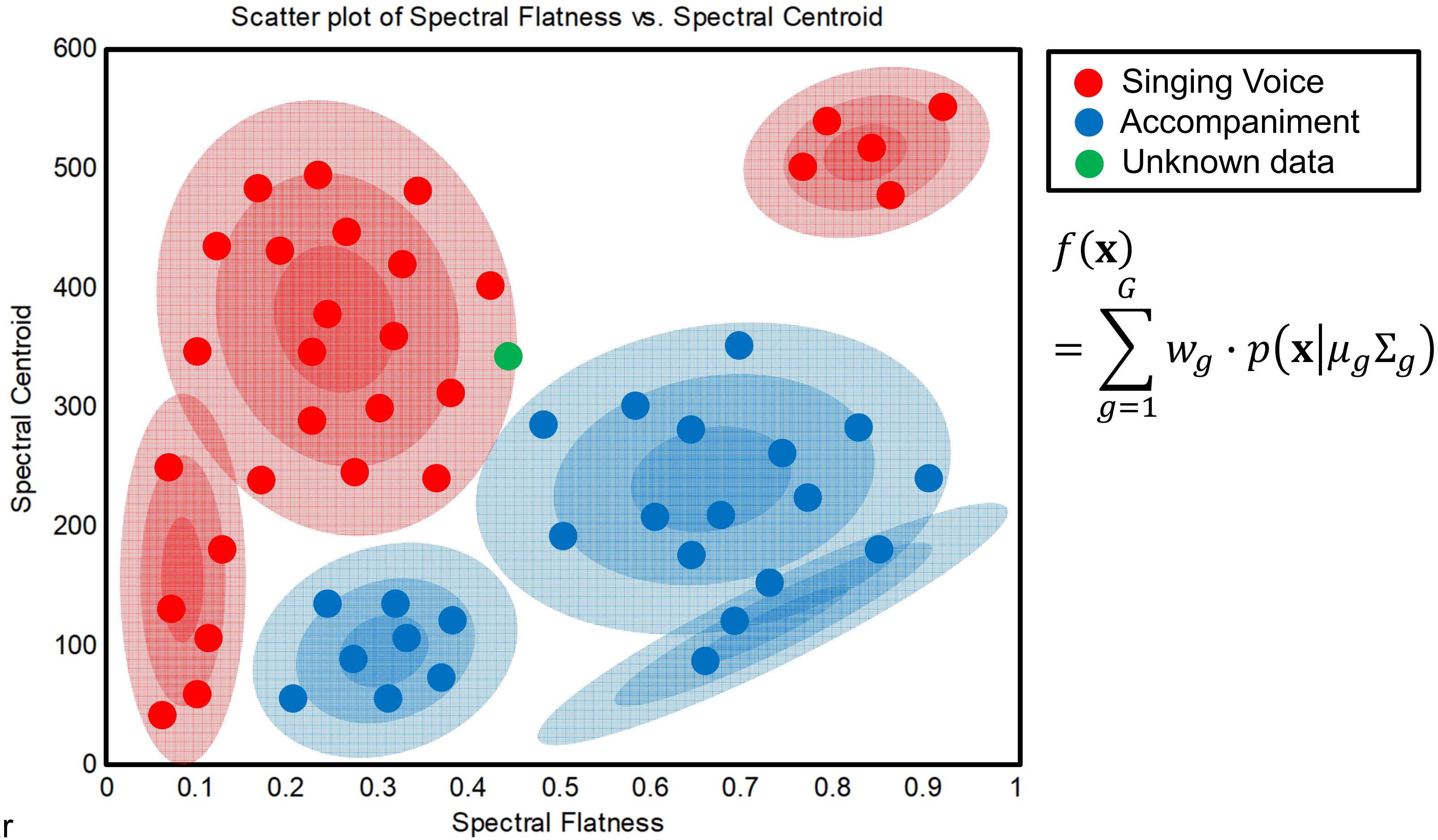
- Random Forests (RF)



Slides:
Christian Dittmar

Classification Methods

- Gaussian Mixture Models (GMM)

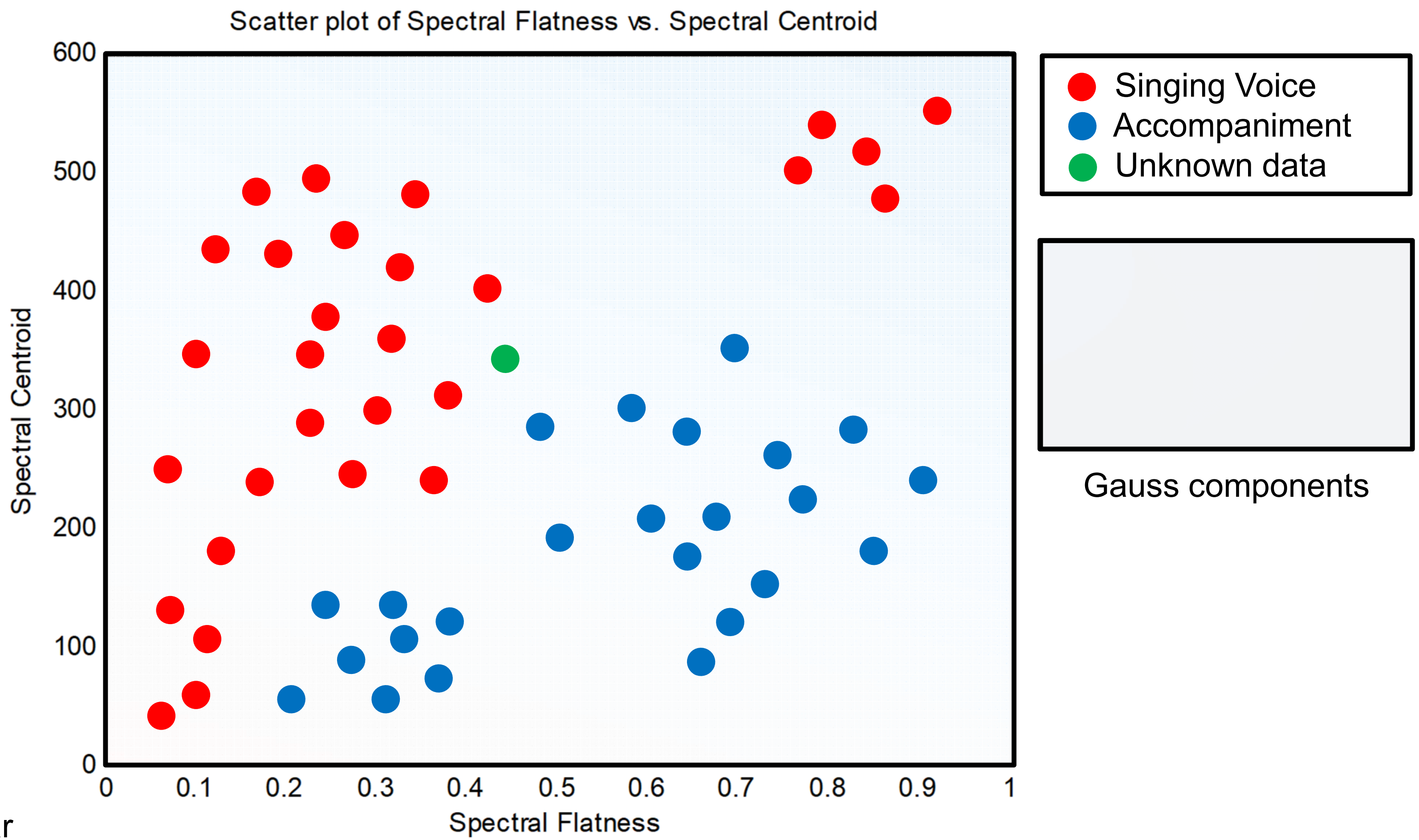


$$f(\mathbf{x}) = \sum_{g=1}^{G} w_g \cdot p(\mathbf{x} | \mu_g \Sigma_g)$$

Slides:
Christian Dittmar

Classification Methods

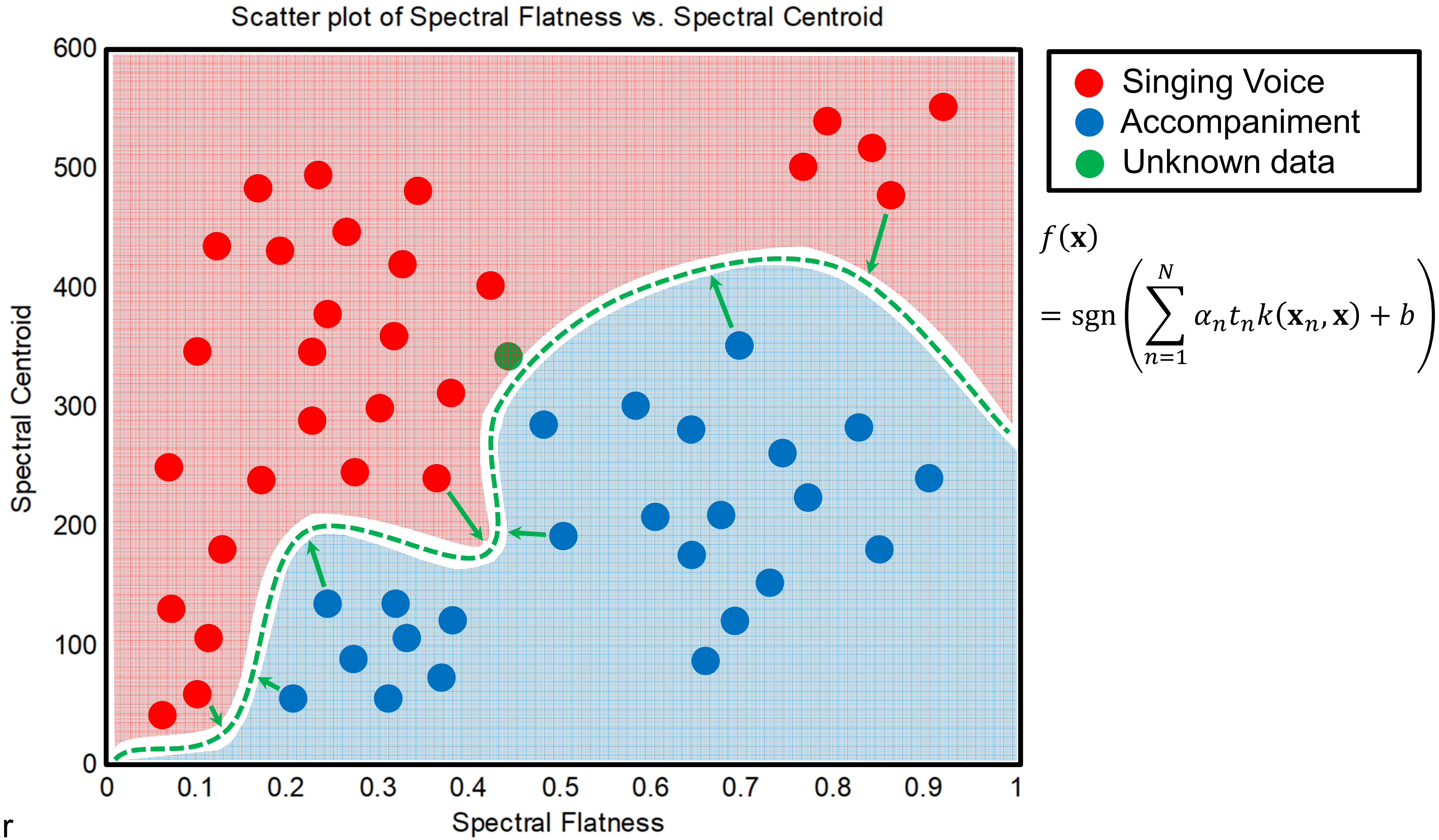
- Gaussian Mixture Models (GMM)



Slides:
Christian Dittmar

Classification Methods

- Support Vector Machines (SVM)

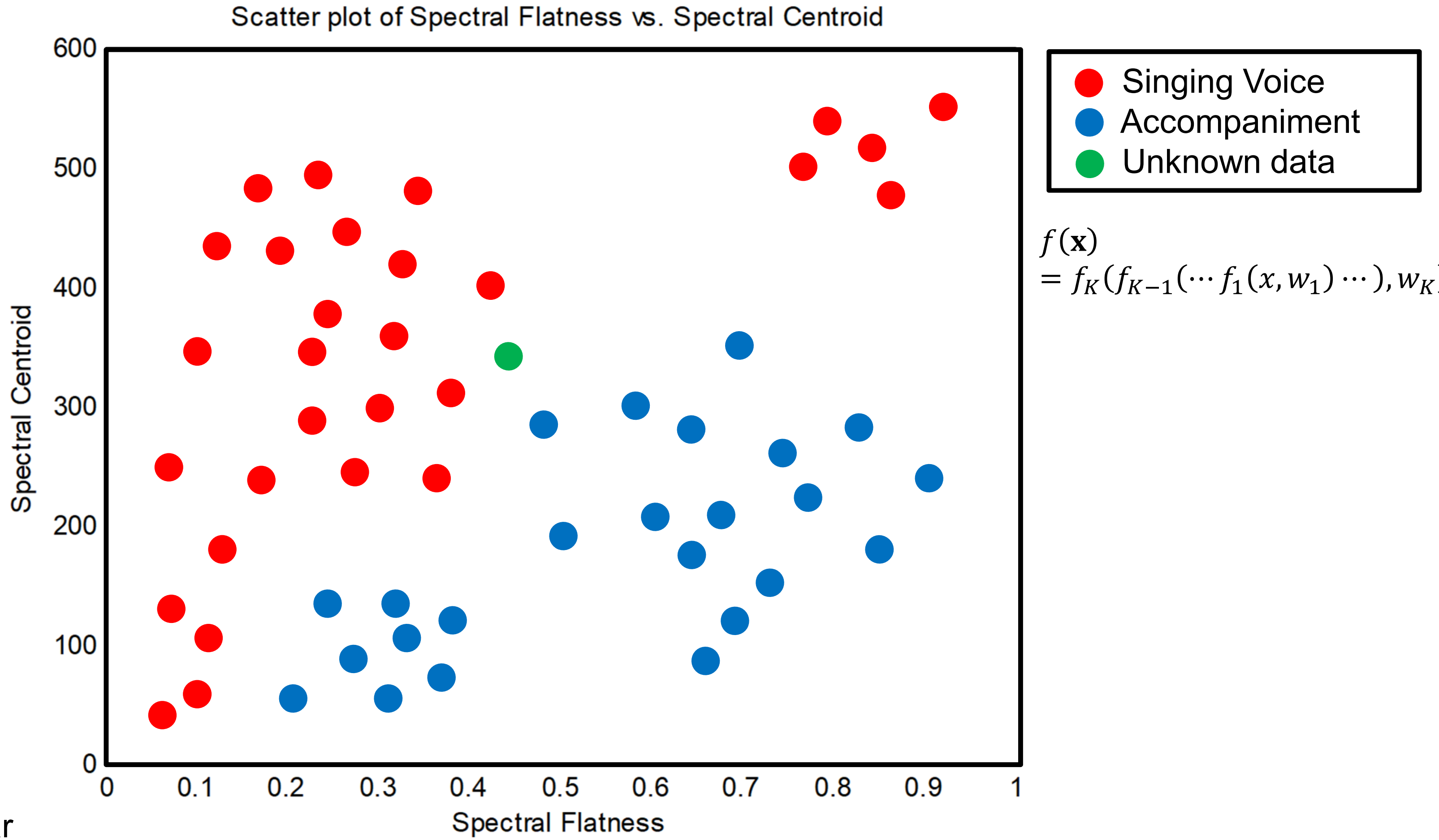


$$f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{n=1}^{N} \alpha_n t_n k(\mathbf{x}_n, \mathbf{x}) + b\right)$$

Slides:
Christian Dittmar

Classification Methods

- Deep Neural Networks (DNN)

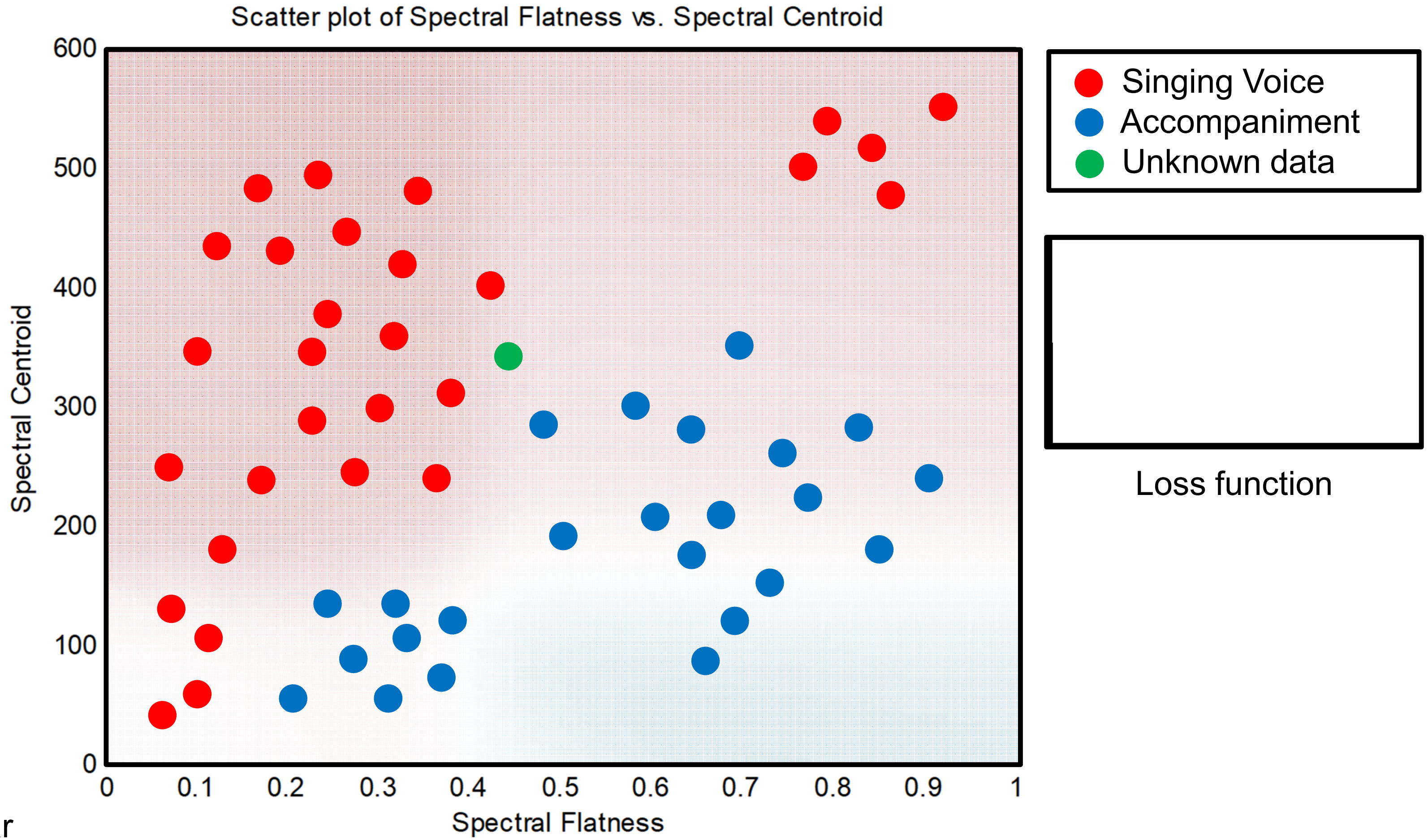


$$f(\mathbf{x}) = f_K(f_{K-1}(\cdots f_1(x, w_1)\cdots), w_K)$$

Slides:
Christian Dittmar

Classification Methods

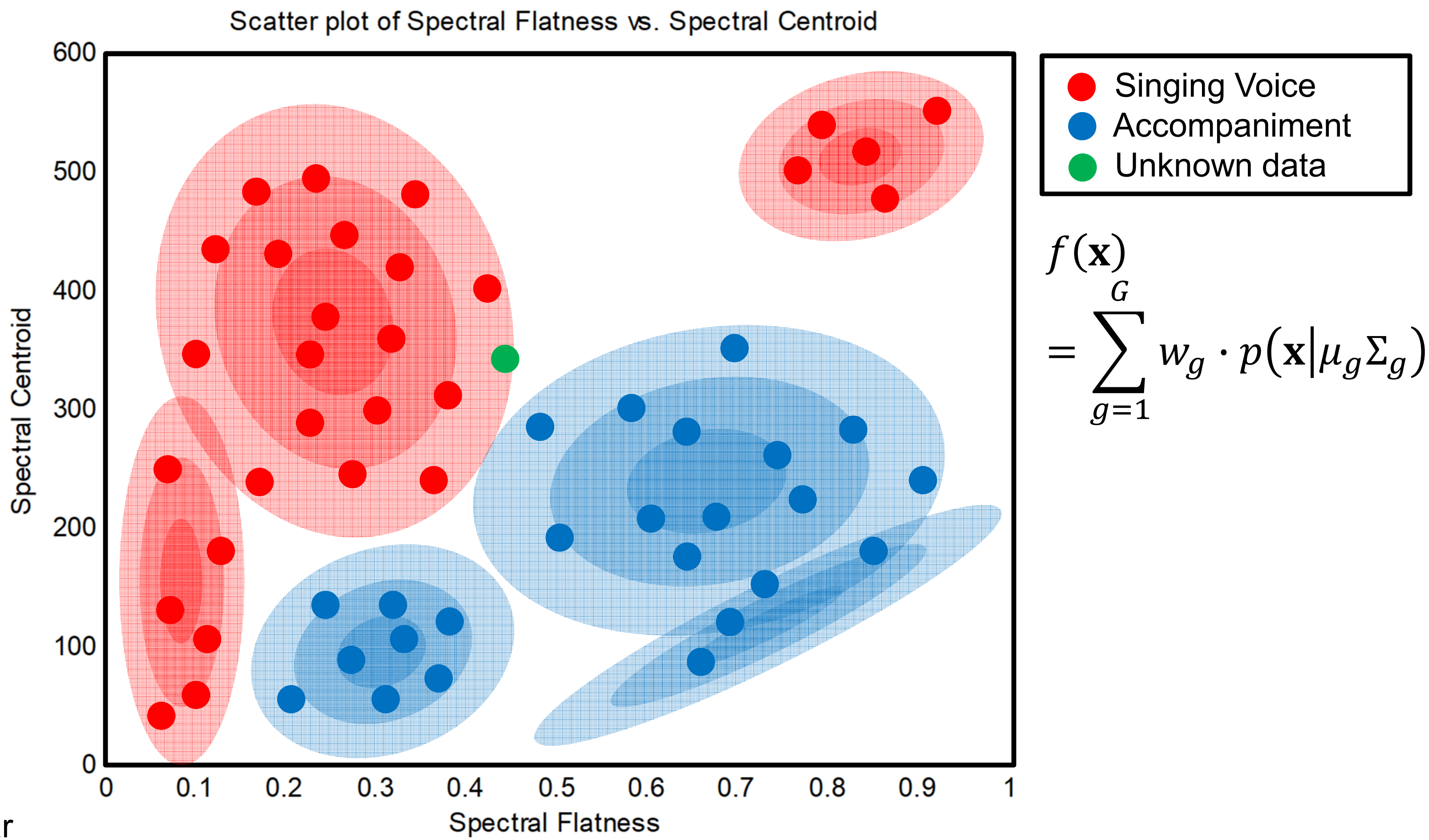
- Deep Neural Networks (DNN)



Slides:
Christian Dittmar

Classification Methods

- Gaussian Mixture Models (GMM)



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

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 %**

Classification Results

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

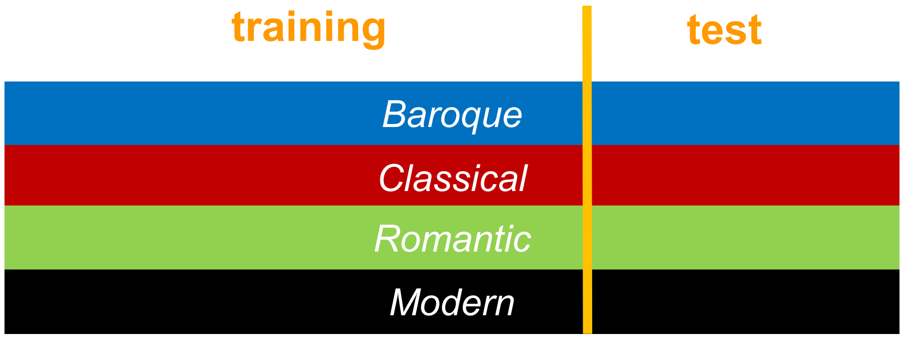
	Full Dataset	**Piano**	**Orchestra**
Standard features	87 %	88 %	85 %
Tonal features	84 %	84 %	86 %
Combined	**92 %**	**86 %**	**80 %**

Overfitting???

Classification Results: Album Effect

- 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 %**



Classification Results: Album Effect

- 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 %**

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: Unseen Data

- Training on **piano**, evaluating on **orchestra** → mean acurracy **65 %**
- Training on **orchestra**, evaluating on **piano** → mean acurracy **64 %**

- Evaluation on **completely unseen data** (composer dataset)
  - Ignoring Beethoven & Schubert
  - Mean accuracy **62.3 %**

Classified Era	Baroque	Classical	Romantic	Modern
Bach	68	5	9	18
Handel	56	23	15	6
Rameau	69	22	6	3
Haydn	25	53	19	3
Mozart	28	51	7	14
Beethoven	16	37	38	9
Schubert	7	16	24	53
Mendelssohn	15	19	55	11
Brahms	6	13	69	12
Dvořak	14	17	65	4
Shostakovich	15	2	8	75

Classification Results: Error Examples

- 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

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