

Music Similarity Analysis

Using the Big Data
Framework Spark

Johannes Schoder

1. November 2019

Fakultät für Mathematik und Informatik
Friedrich-Schiller-Universität Jena



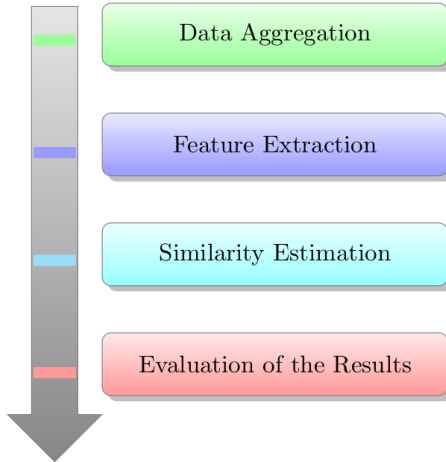
- Music Streaming Services - Datenmengen wachsen
 - Amazon Prime Music: 2 Millionen
 - Spotify: 35 Millionen
 - Soundcloud+: 150 Millionen

Gibt es das nicht schon?

- "Collaborative Filtering"
 - ...
- Forschung
 - Viele Einzelaspekte separat beleuchtet (Melodie, Rhythmus oder Klangfarbe)
- Was fehlt?
 - Parametrisierbare Empfehlungen (Schwerpunkte legen)

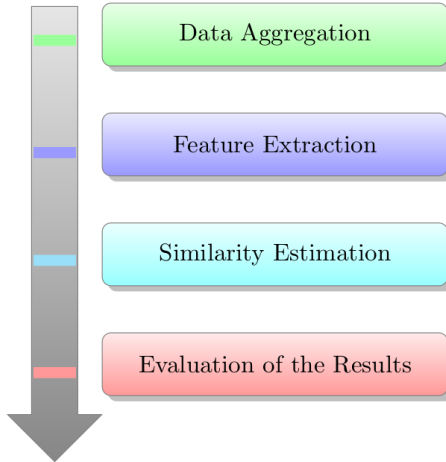
Datenextraktion Roadmap

■ Aufgaben



- Free Music Archive - ca. 103.000 Lieder
 - Soundbeispiel
- Private Musiksammlung - ca. 8000 Lieder
- 1517 Artists Dataset - 3180 Lieder
- Covers80 - 164 Lieder
- Summe: 1TB - ca 114000 verwertbare Songs

■ Data Aggregation - Check



Feature Extraction Roadmap

- Verschiedene Features parallel auf dem ARA-Cluster aus den Daten extrahieren:
 - Melodie
 - Noten
 - Chromagramm
 - Rhythmus
 - Beat Histogram
 - Rhythm Histogram
 - Rhythm Pattern
 - Klangfarbe
 - MFCCs

Ideale Features?

■ Melody Extraction



Abbildung: Für Elise [1]

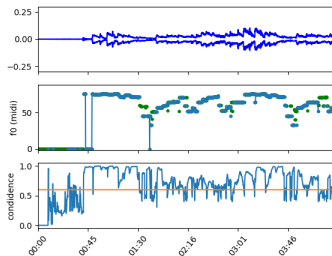


Abbildung: Pitch (Aubio)

Nicht ganz so ideale Features

■ MIDI Noten aus Pitch vs. Original



Abbildung: Für Elise [1]



Abbildung: Erkannte Melodie (Aubio)

- Play: Original
- Play: Noten
- Play: Erkannt

■ E-Gitarre (verzerrt) vs. Klavier

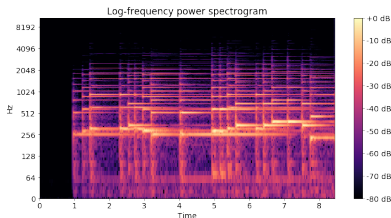


Abbildung: Klavier

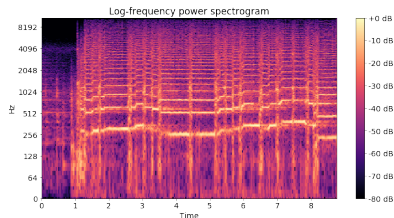


Abbildung: Gitarre

■ Play: Klavier

■ Play: Gitarre

Melodie Bandpass

■ 128 Hz bis 4096 Hz

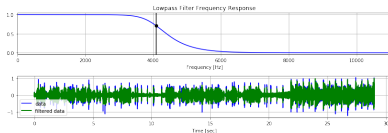


Abbildung: Tiefpass 128 Hz

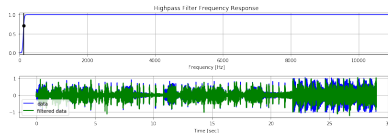
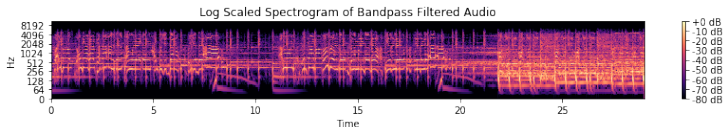


Abbildung: Hochpass 4096 Hz



Melodieextraktion

■ Noten aus Chromagramm (Oberwellen minimieren)

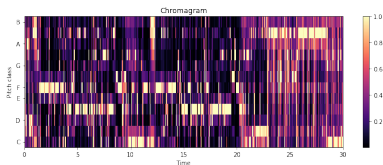


Abbildung: Sia - Chandelier - Chromagramm

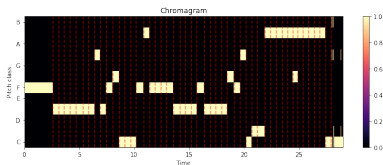


Abbildung: Sia - Chandelier - "Noten"

- Noten zu String: [A, A, B, B, B, D, ... , D]
- Strings vergleichen (Levenshtein-Distanz) für Ähnlichkeitsschätzung

Kreuzkorrelation zwischen Chromagrammen

- Kreuzkorrelation zwischen Matrix X ($P \times Q$ Matrix) und Y ($M \times N$ Matrix) zu Vektor C

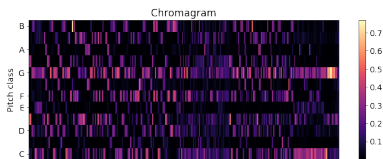
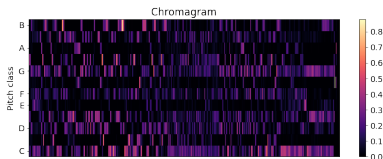
$$C(l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(m, n) Y(m, n - l) \quad (1)$$

$$0 \leq l \leq N - Q \quad (2)$$

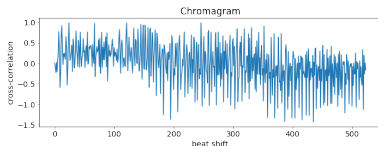
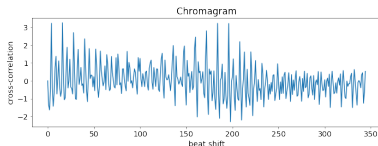
- Ergebnisvektor wird Hochpassgefiltert

Kreuzkorrelation zwischen Chromagrammen

■ Zwei am Takt ausgerichtete Chromagramme

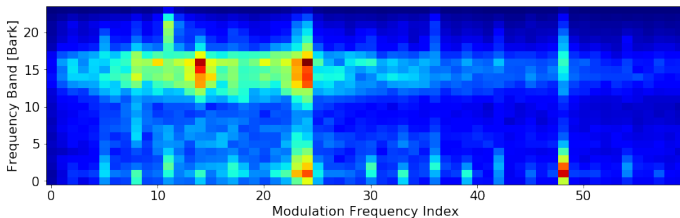


■ Ergebnisvektor (HP gefiltert; Cover vs. verschiedene Lieder)

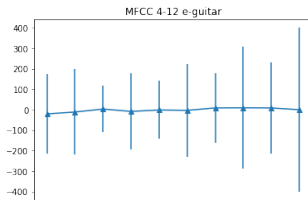
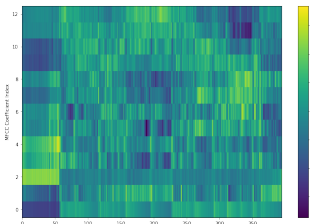


Und das waren nur die Melodiefeatures...

■ Rhythmus (z.B. Rhythm Patterns)



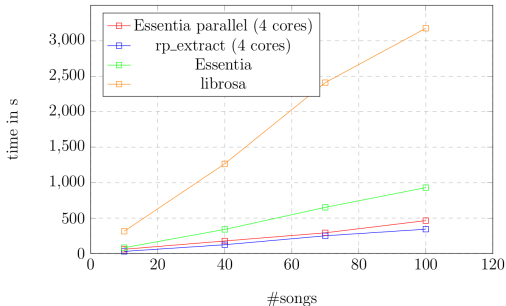
■ Klangfarbe/ Timbre (MFCCs)



Feature Extraction Timing

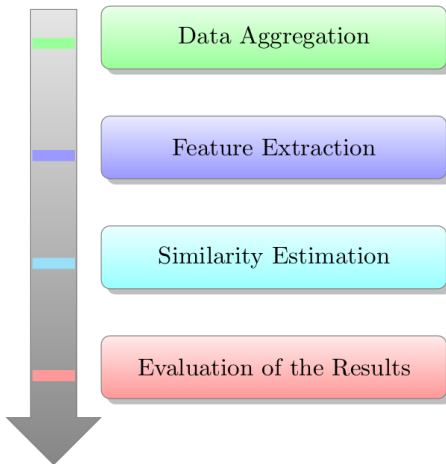
■ Feature Extraktion mit MPI4PY

- ca 5 min für 100 Lieder auf 4 CPU Cores



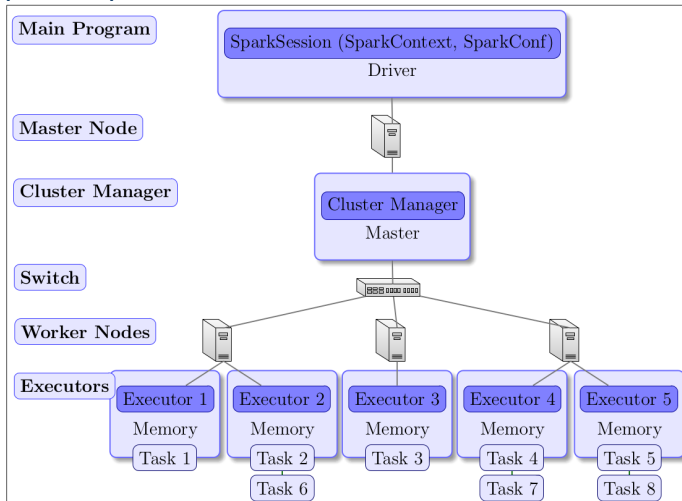
- ca. 32,5 min für 102.000 Lieder auf 36 * 18 CPU Cores mit je 10GB RAM/Core

■ Feature Extraction - Check



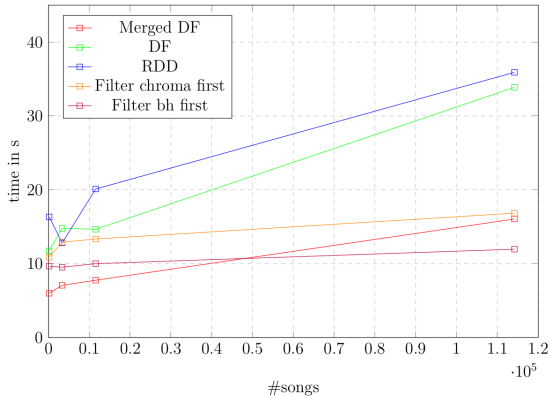
- Melodie
 - Levenshtein-Distanz zwischen zwei Strings
 - Kreuzkorrelation zwischen zwei Chromagrammen
- Klangfarbe
 - Jensen-Shannon Divergenz
 - Kullback-Leibler Divergenz
 - Euklidischer Abstand
- Rhythmus
 - Euklidischer Abstand (RP, RH, BH)

■ Apache Spark Cluster

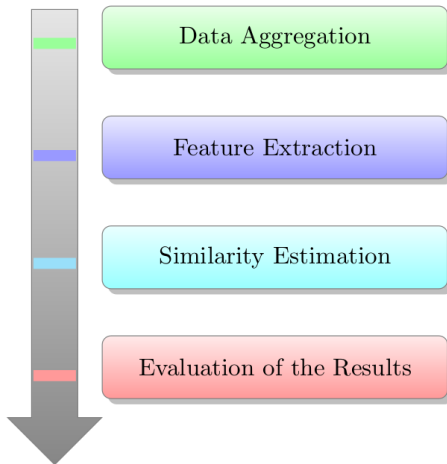


■ Ähnlichkeitsschätzung mit Spark

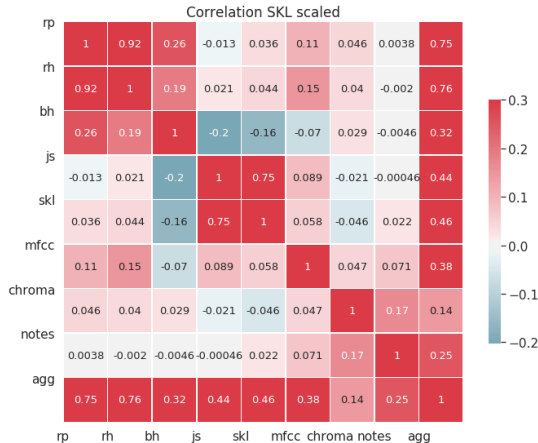
■ 12 sec für 114.000 Lieder



■ Nächster Schritt: Evaluation

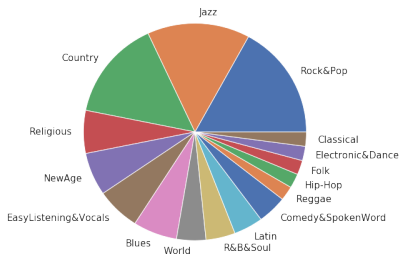
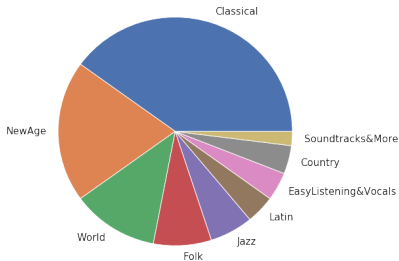


Korrelation der verschiedenen Distanzen

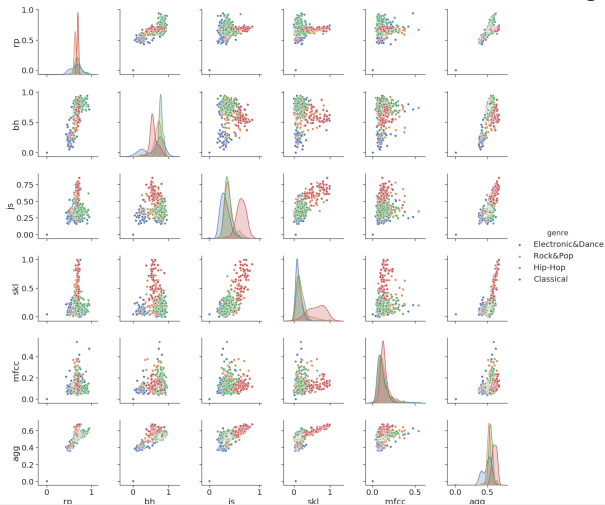


Genre Recall Rate

- Empfehlungen für 10 Classical und 10 Rock/Pop Songs basierend auf Rhythmus- und Timbre-Features (3180 Testsongs)



■ Distanzen aller Lieder zu einem Electronic-Song (3180 Testsongs)



- Song request: 100 Meisterwerke der Klassik - Mozart - Alla Turca (Allegretto) (private collection)
 - Mozart Collection / CD31 / KV331-3 Alla turca allegretto (private collection)
 - Piano Collection / CD25 - Mozart - Alla Turca Allegretto (private collection)
 - Piano Perlen / Mozart - Türkischer Marsch (private collection)
 - FRITZ STEINEGGER - RONDO ALLA TURCA KV 331 (1517-Artists)
 - 136071 (2Kutup - We Shall Cuddle Up And Sleep) (FMA dataset)
 - Sean Bennett - Variations on the Turkish March (1517-Artists)

Cover Song Recognition Rate - covers80

features	detected covers
chroma	30
chroma + notes	27
chroma + skl	26
chroma + notes + rp	24
chroma + rp	22
chroma + skl + rp	22
chroma + mfcc	19
chroma + js + rp	17
chroma + js	17
notes	17
chroma + mfcc + rp	15
all	15
notes + rp	13
mfcc + notes + rp	7
rp	7
mfcc + js + skl	3

features	detected covers
chroma	33
chroma + notes	31
chroma + notes + rp	30
chroma + skl	29
chroma + rp	29
chroma + skl + rp	26
chroma + mfcc + rp	24
notes	23
all	23
chroma + mfcc	22
chroma + js + rp	22
chroma + js	21
notes + rp	19
rp	15
mfcc + notes + rp	14
mfcc + js + skl	10

Table 5.1: Cover recognition rate - Top 1 Table 5.2: Cover recognition rate - Top 5

- Song Request

- ...

- MIREX 2020
 - Paper?
- Standalone Version?
- Bugfixes/ Erweiterungen (skalieren)/ Verbesserungen möglich - stehen in der Arbeit



“Für elise.” [Online]. Available: https://upload.wikimedia.org/wikipedia/commons/a/a9/BH_116_Vergleich.png

Danke für die Aufmerksamkeit

- Fragen?

- ...