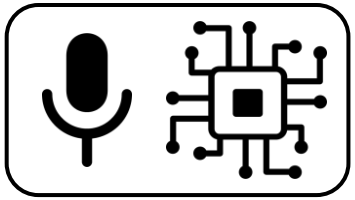
Computational Analysis of Sound and Music

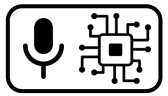


Music Information Retrieval – Harmonic Analysis

Dr.-Ing. Jakob Abeßer

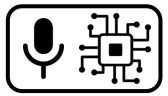
Fraunhofer IDMT

jakob.abesser@idmt.fraunhofer.de



Outline

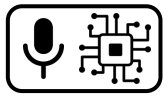
- **Fundamental concepts**
- Chromagram
- Chord Recognition



Fundamental Concepts

- Western music is based on **equal-temperament tuning**
  - Each octave is divided into 12 semitones

- Human **pitch perception** is **periodic**
  - 2 pitches one octave apart are perceived as similar



Fundamental Concepts

Chroma

- **Pitch** ⟶ Chroma + Tone height
  - Chroma: C, C#, D, D#, …, B (12 classes)
  - Tone height: Octave number

Figure 3.3a from [Müller, FMP, Springer 2015]

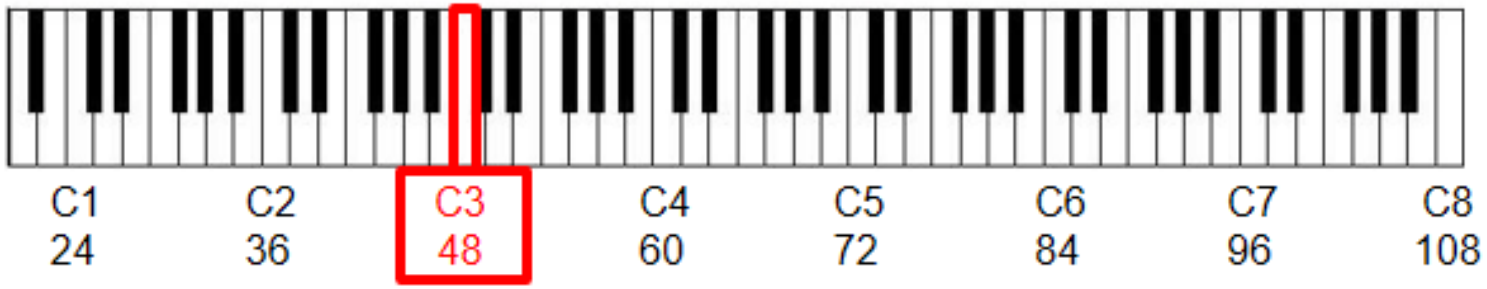
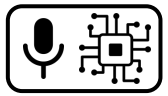


Fig-M2-1

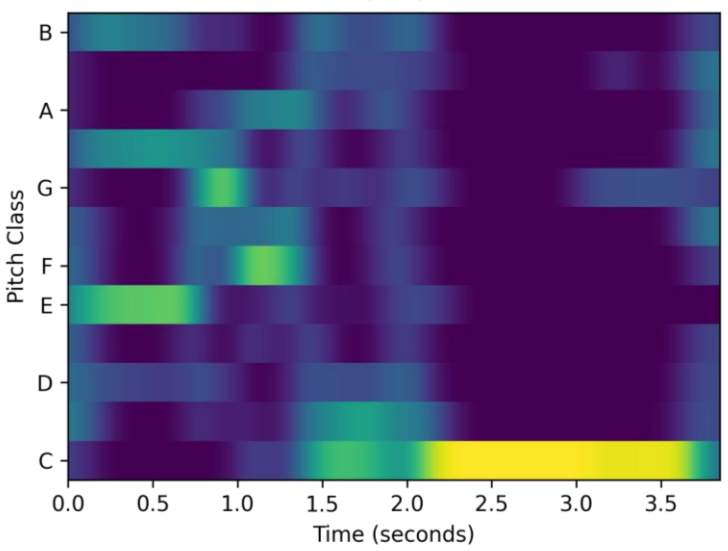
Outline

- Fundamental concepts
- **Chromagram**
- Chord Recognition



Chromagram

- Intensity of pitch classes over time (similar to spectrogram)



Own

Aud-A2-2

- Applications
  - Analyze and compare harmonic and melodic content in music
  - Main audio representation for chord & key recognition

Chromagram

- Basic processing pipeline
  - Pitch features derived from STFT or CQT
  - Summing up over pitch sub-bands across octaves to pitch class intensities

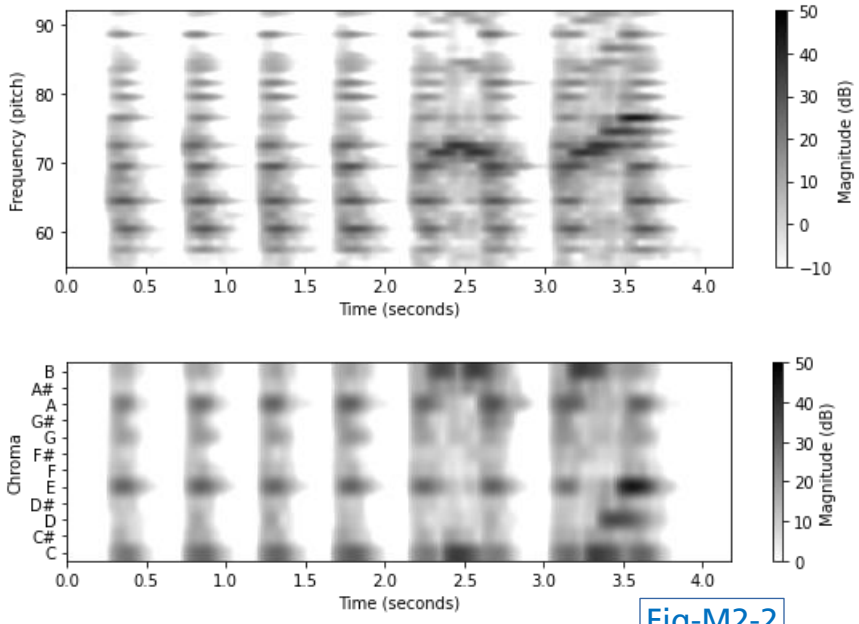


Fig-M2-2

Pitch Feature

Chromagram

`librosa.feature.chroma_stft`

`librosa.feature.chroma_cqt`

Chromagram

Chromagram Variants

- Chroma Energy Normalized Statistics (CENS) [Müller et al., 2005]
  - Derived from Constant-Q chromagram
  - Post-processing steps → invariance to timbre and dynamics
    - L1 normalization per chroma vector
    - Amplitude quantization
    - Smoothing with sliding window & down-sampling

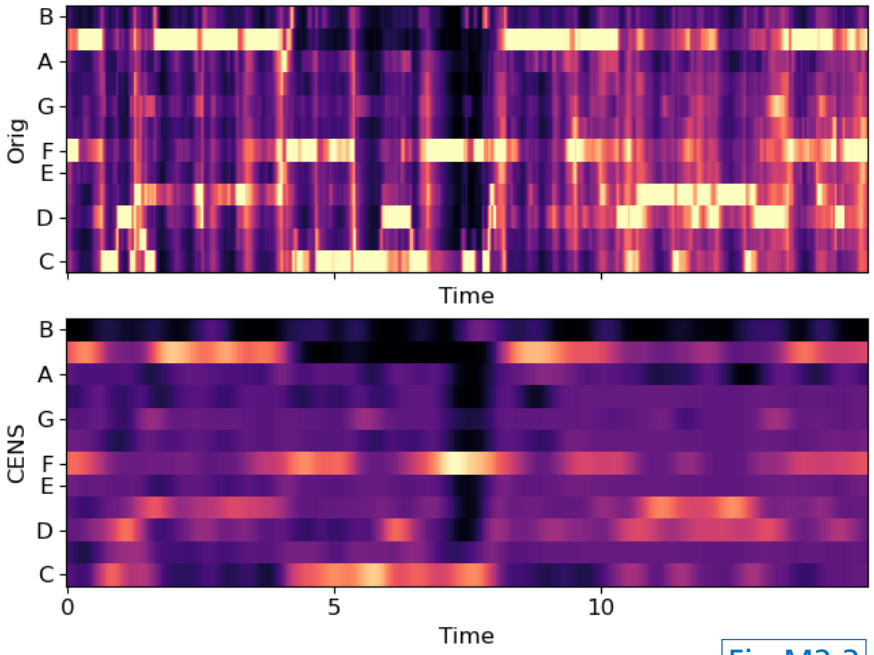


Fig-M2-3

`librosa.feature.chroma_cens`

Chromagram

Chromagram Variants

- Deep Chroma [Korzeniowski & Widmer, 2016]
  - Input: quarter-tone resolution
  - MLP (3 hidden layers, 512 units)

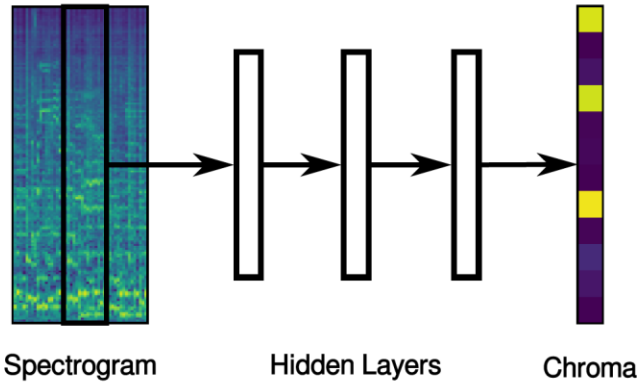


Fig-M2-4

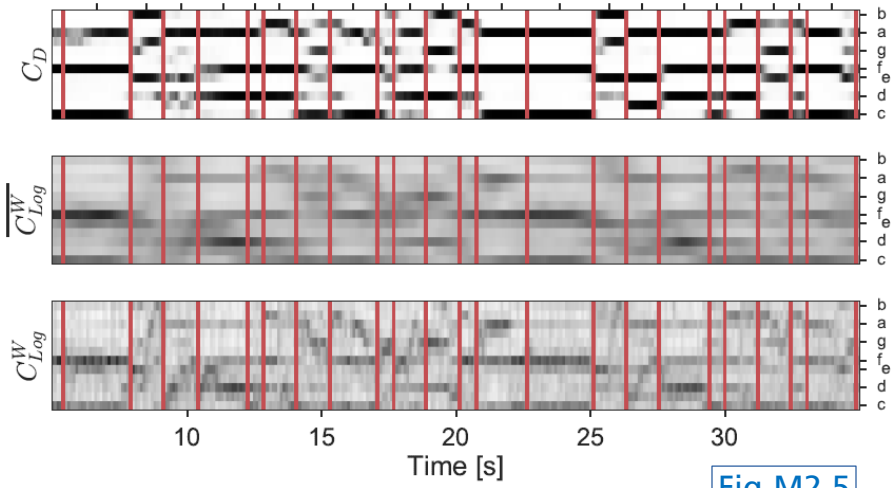
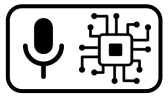


Fig-M2-5

`madmom.features.chords`

Outline

- Fundamental concepts
- Chromagram
- **Chord Recognition**



Chord Recognition

Basic Principles

- Chroma Feature Representations → Chord classes

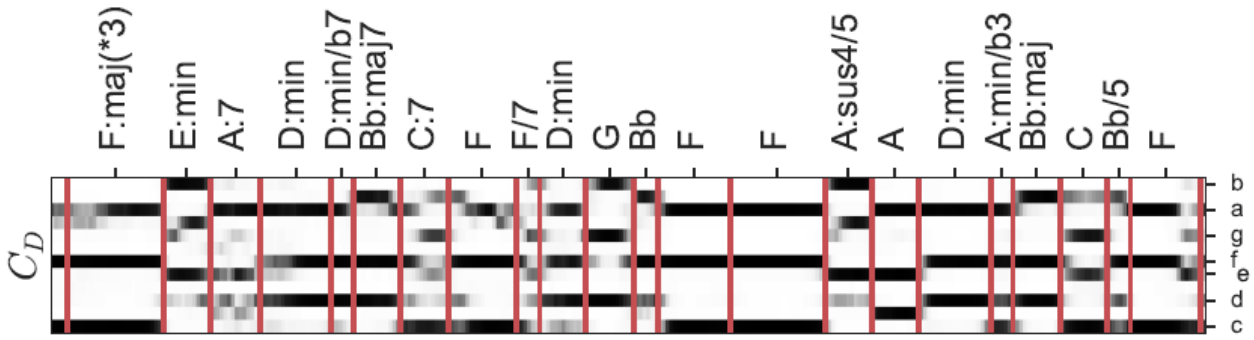


Fig-M2-5

- Temporal Segmentation into short, overlapping frames to capture local chords / chord changes

Chord Recognition

Challenges

- Many chords share chord tones → ambiguities

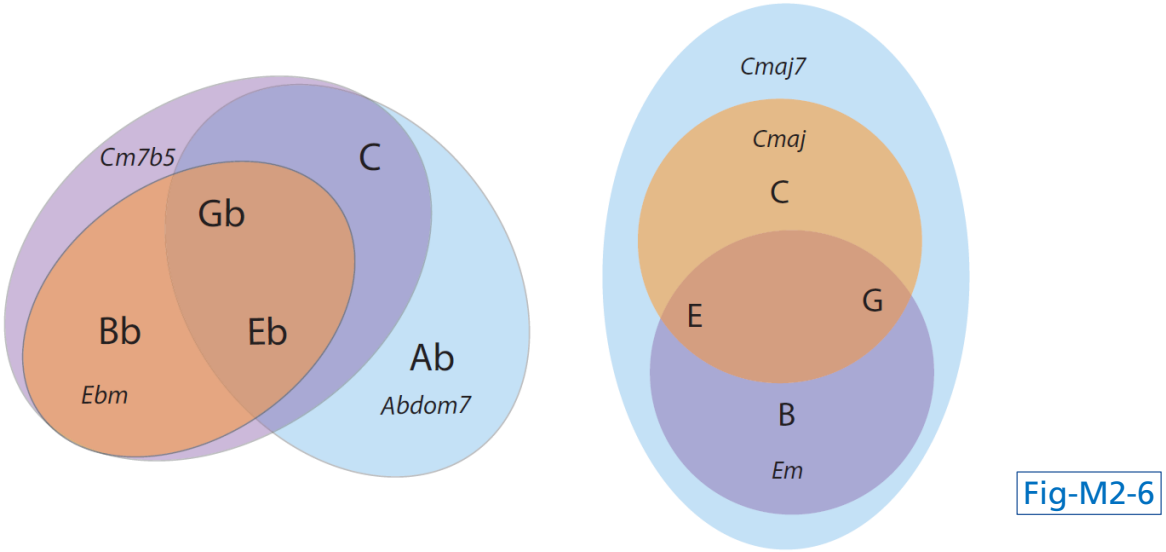


Fig-M2-6

- Chord inversions (C, C/E, C/G) → same pitch classes, different “pitch ordering”
  - Separate root note & bass note pitch class classification

Chord Recognition

Traditional Approach: Template Matching

- Chord tones → Binary templates $T \in \mathbb{R}^{12}$
- Brute-force comparison between chromagram frame and all chord templates
  - Cross-correlation, cosine similarity …

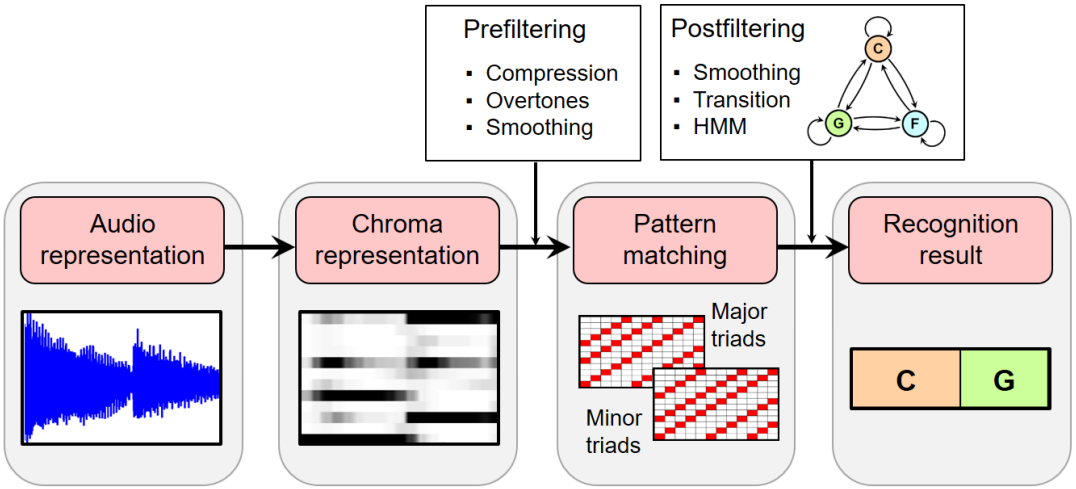
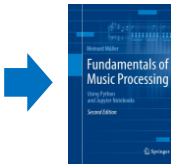


Figure 5.13 from [Müller, FMP, Springer 2015]

Fig-M2-7



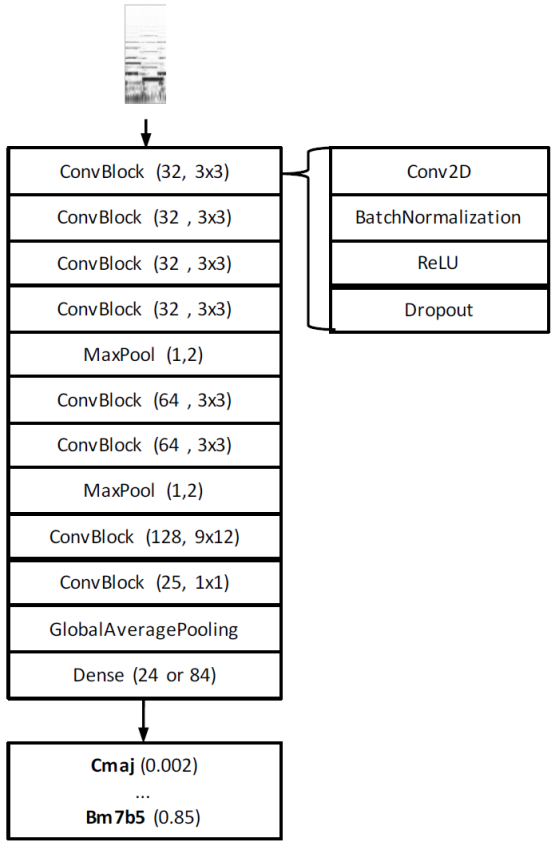
FMP Notebooks

Chord Recognition

DL-based Method

- [Nadar et al., 2019]
  - STFT + log frequency axis (24 bins per octave)
  - Spectrogram patches: 1.5 s duration
  - 84 chord classes (7 chord type)
    - Three-voiced chords (maj, min)
    - Four-voiced chords (maj7, min7, dom7, m7b5) + “powerchord”: 1+5

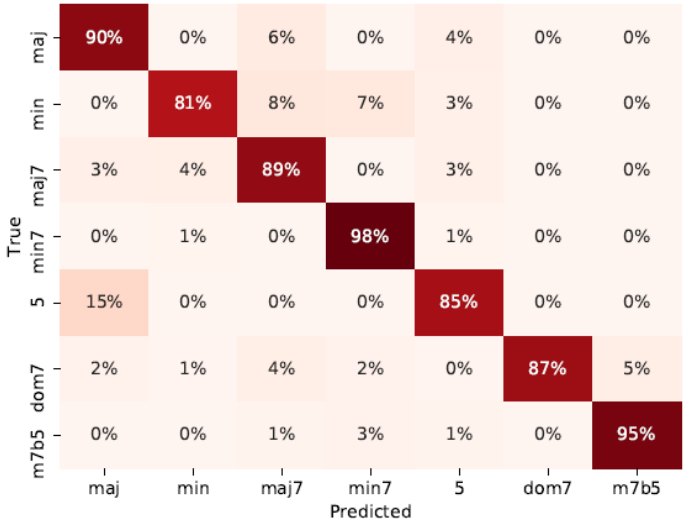
- Give 1 example method



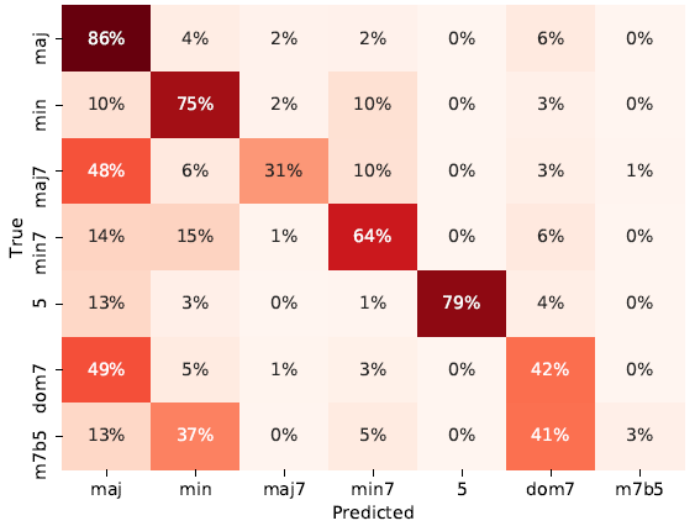
Chord Recognition

Machine Learning Models

- Isolated chord recordings



- Mixed chord recordings



https://machinelistening.github.io/casm

Programming session

Fig-A2-13

References

Images

Fig-M2-1: [Müller, 2015], Fig. 3.3a

Fig-M2-2: https://www.audiolabs-erlangen.de/resources/MIR/FMP/C3/C3S1_SpecLogFreq-Chromagram.html

Fig-M2-3: https://librosa.org/doc/main/_images/sphx_glr_plot_chroma_007.png

Fig-M2-4: [Korzeniowski & Widmer, 2016], p. 39, Fig. 1

Fig-M2-5: [Korzeniowski & Widmer, 2016], p. 42, Fig. 7

Fig-M2-6: [Nadar et al., 2019], p. 553, Fig. 4

Fig-M2-7: [Müller, 2015], Fig. 5.13

References

References

Müller, M. (2021). Fundamentals of Music Processing - Using Python and Jupyter Notebooks (2nd ed.). Springer.

Müller, M. & Ewert, S. (2011). Chroma Toolbox: MATLAB Implementations for Extracting Variants of Chroma-Based Audio Features. Proceedings of the International Conference on Music Information Retrieval (ISMIR), 2011.

Müller, M. & Ewert, S. (2010). Towards timbre-invariant audio features for harmony-based music. IEEE Transactions on Audio, Speech, and Language Processing, vol. 18, no. 3, pp. 649–662, 2010

Müller, M., Kurth, F., & Clausen, M. (2005). Audio matching via chroma-based statistical features. Proceedings of the International Conference on Music Information Retrieval (ISMIR), pp. 288-295, 2005.

Korzeniowski, F. & Widmer, G. (2016). Feature Learning for Chord Recognition: The Deep Chroma Extractor. In Proceedings of the 17th International Society for Music Information Retrieval Conference (ISMIR), New York, USA.

Nadar, C.-R., Abeßer, J., & Grollmisch, S. (2019). Towards CNN-based Acoustic Modeling of Seventh Chords for Automatic Chord Recognition. Proceedings of the 16th Sound & Music Computing Conference (SMC), Málaga, Spain