

PAT 463/563 (Fall 2025)

Music & AI

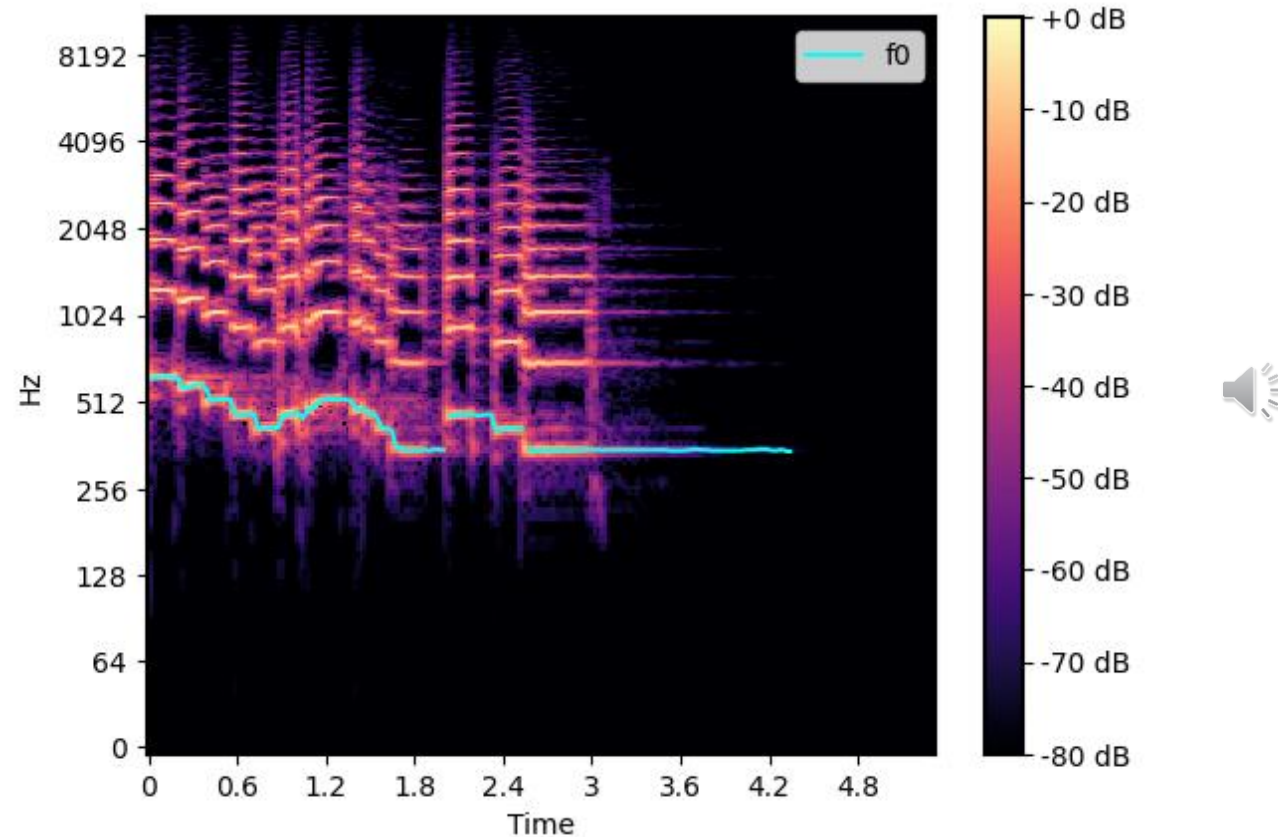
Lecture 10: Music Analysis

Instructor: Hao-Wen Dong

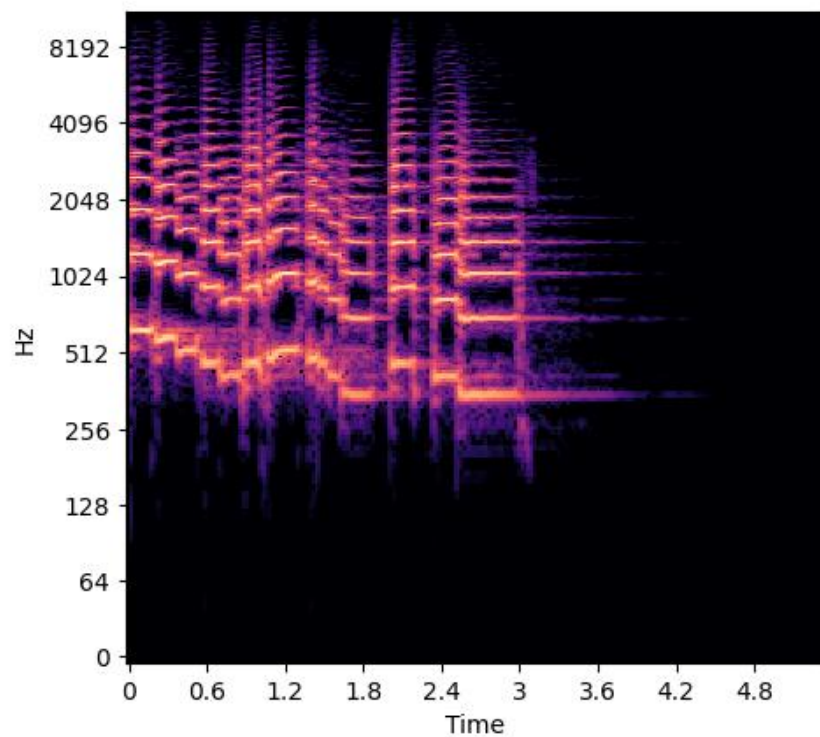
Music Transcription

Fundamental Frequency (F0)

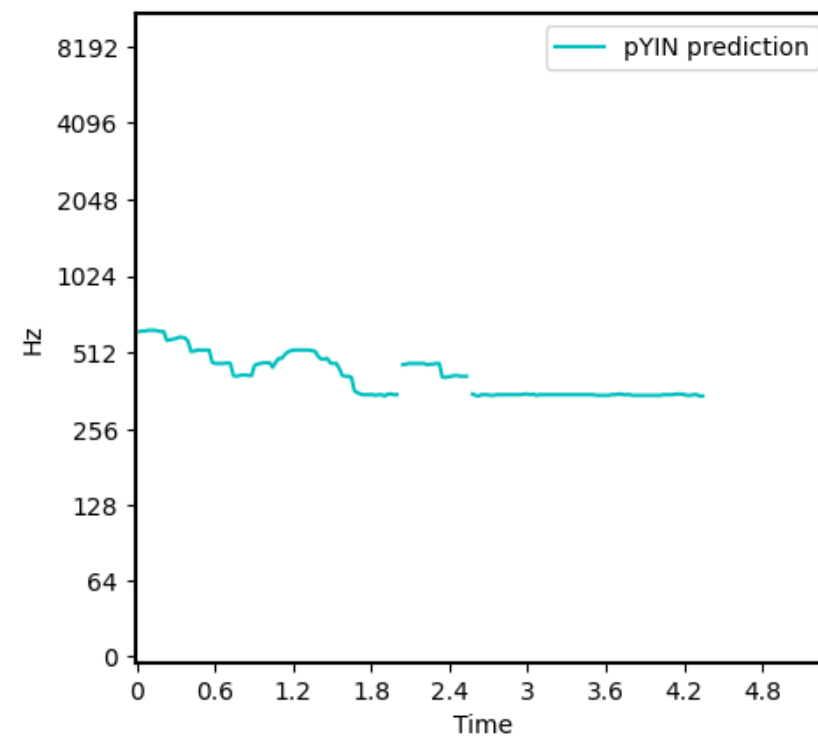
- **Definition:** The **lowest frequency component** of a waveform



Fundamental Frequency (F0) Estimation



→ **F0 Estimation** →



Auto-tune Artist



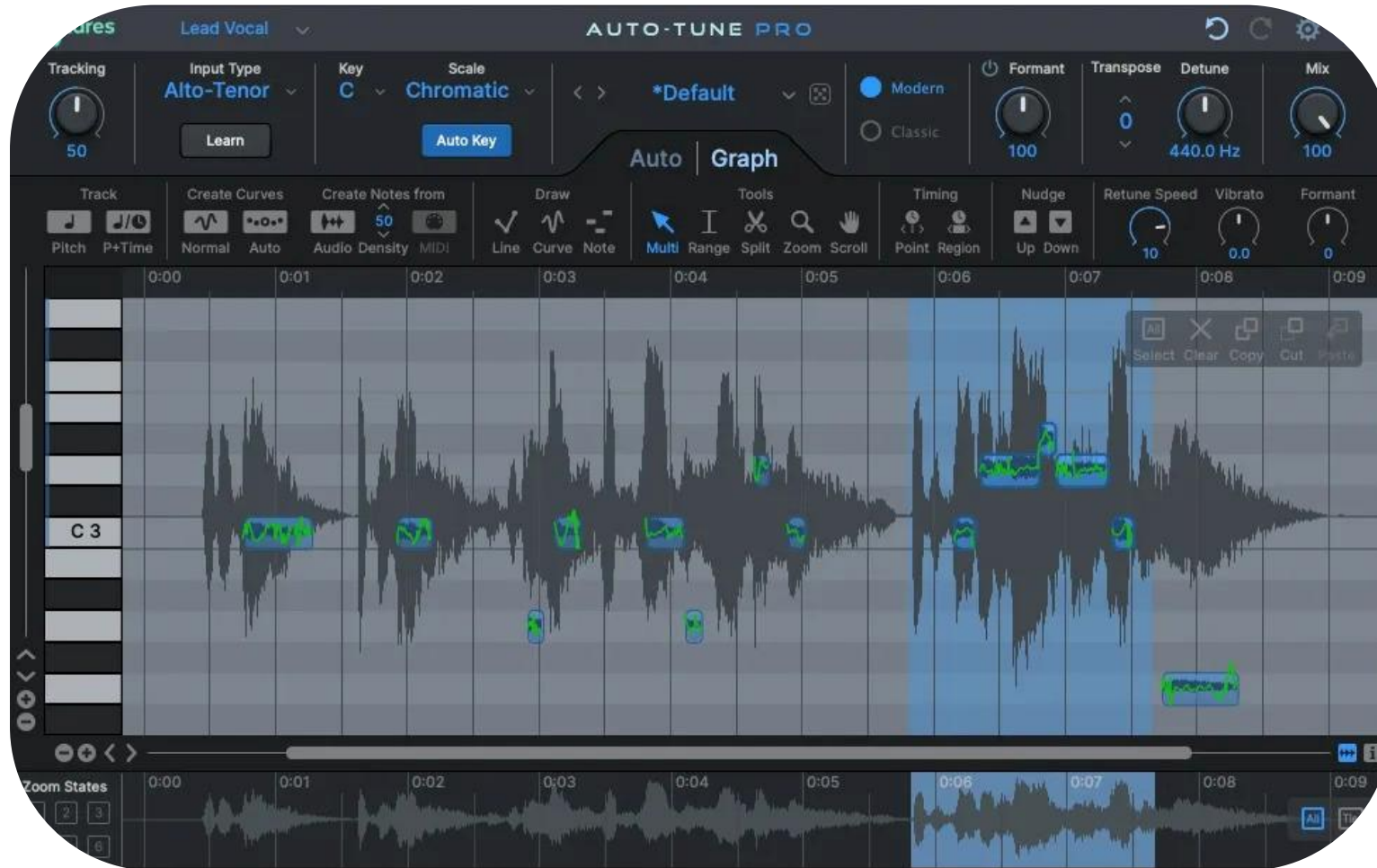
(Source: Antares Audio Technologies)

| Auto-tune



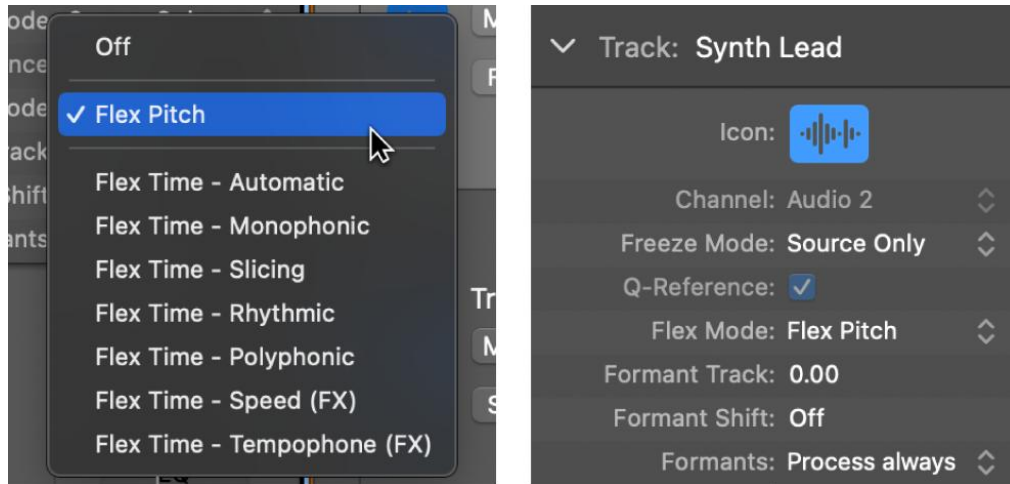
[youtube.com/shorts/
Kg8OSbKRETA](https://youtube.com/shorts/Kg8OSbKRETA)

Auto-tune Pro

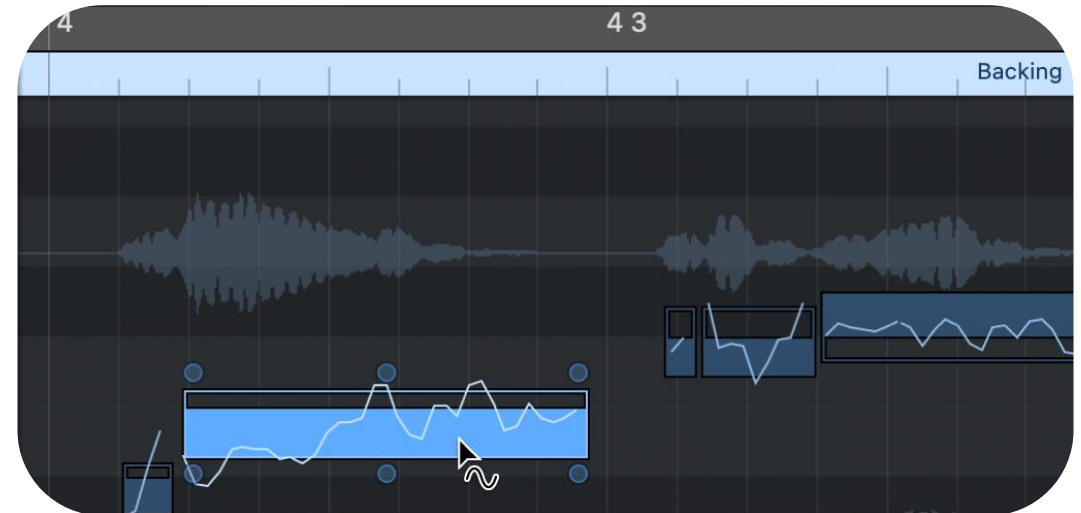


(Source: Antares Audio Technologies)

Pitch Correction in Logic Pro



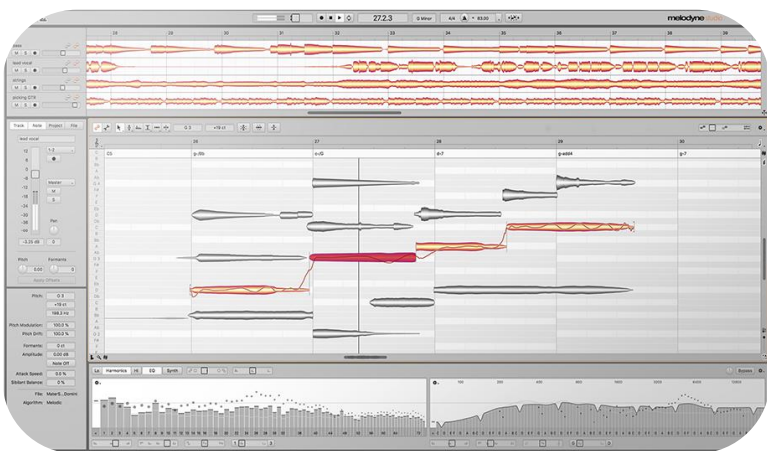
(Source: Logic Pro User Guide)



(Source: Logic Pro User Guide)

Other Auto-tune & Pitch Correction Tools

Melodyne



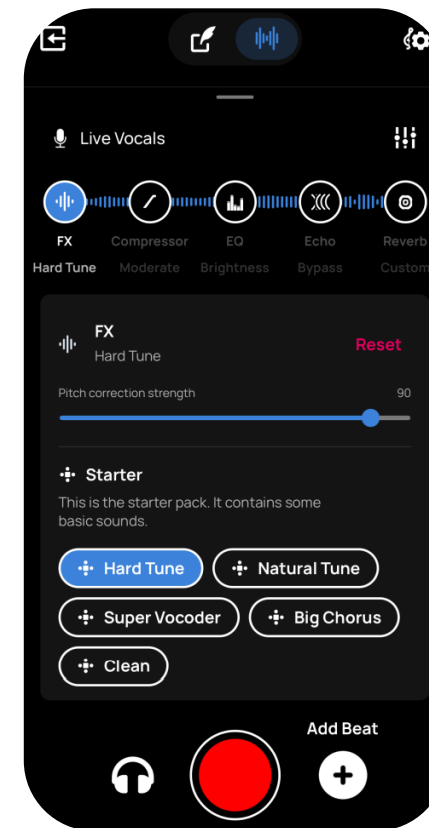
(Source: Celemony)

iZotope Nectar



(Source: iZotope Team)

Voloco



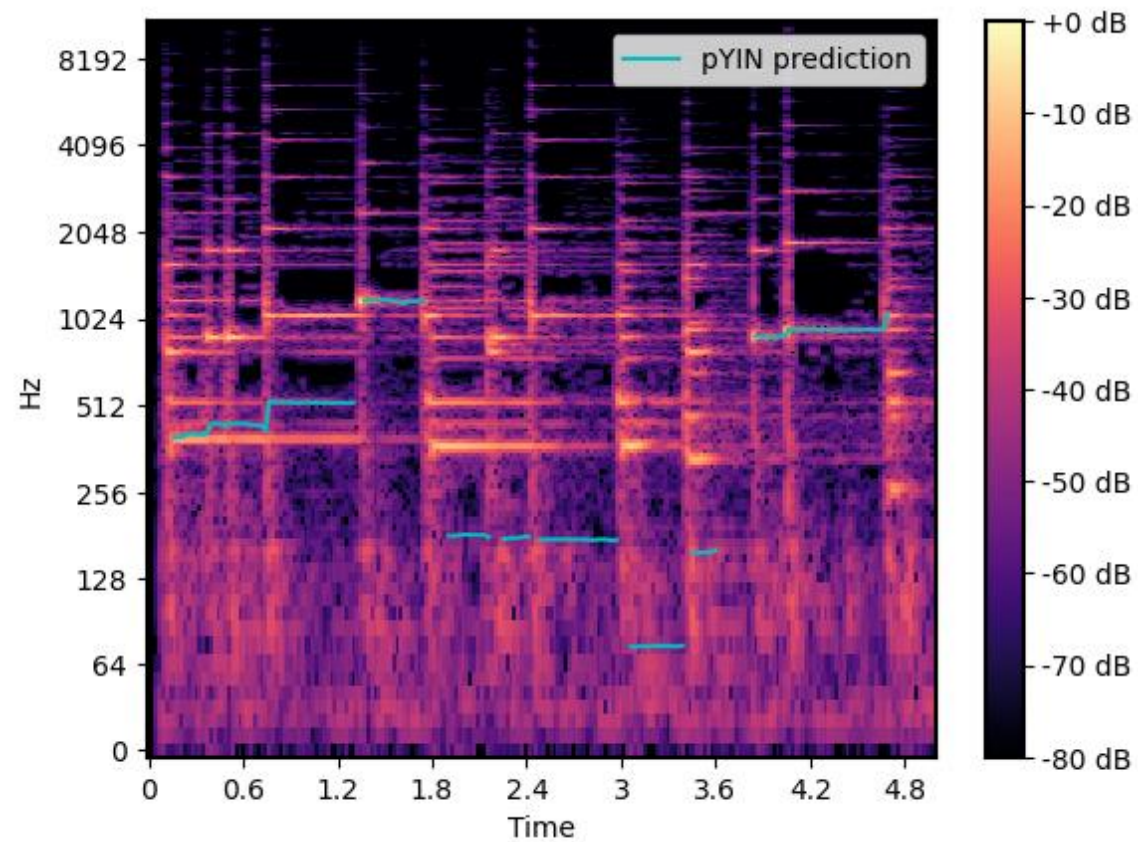
(Source: Voloco User Manual)

celemony.com/en/melodyne/what-is-melodyne
izotope.com/en/learn/why-upgrade-to-nectar-4.html
resonantcavity.com/assets/docs/voloco-user-manual.pdf

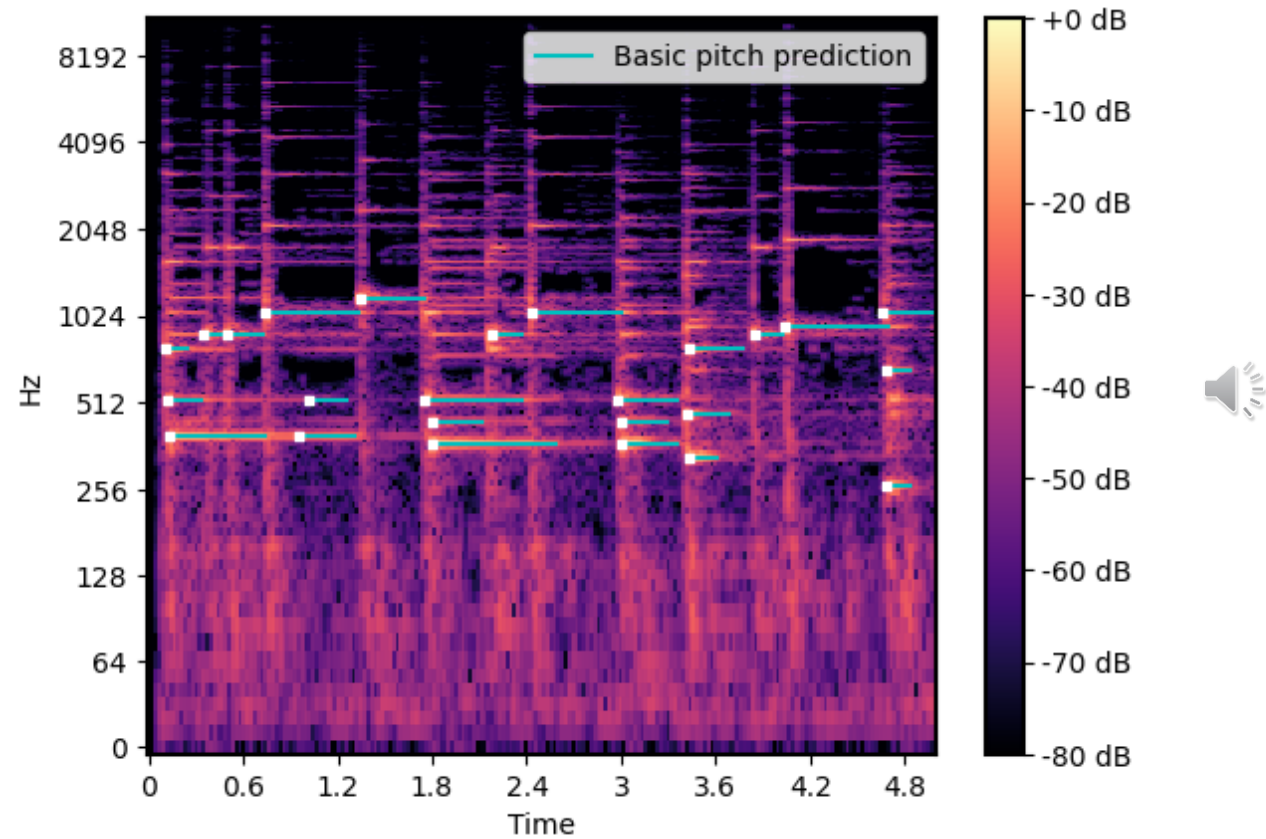
F0 Estimation Models

- **pYIN** (de Cheveigné et al., 2002)
 - `librosa.pyin`
- **CREPE** (Kim et al., 2018)
 - github.com/marl/crepe
- **PESTO** (Riou et al., 2023)
 - github.com/SonyCSLParis/pesto

Polyphonic F0 Estimation



Polyphonic F0 Estimation



basicpitch.spotify.com

Polyphonic F0 Estimation Models

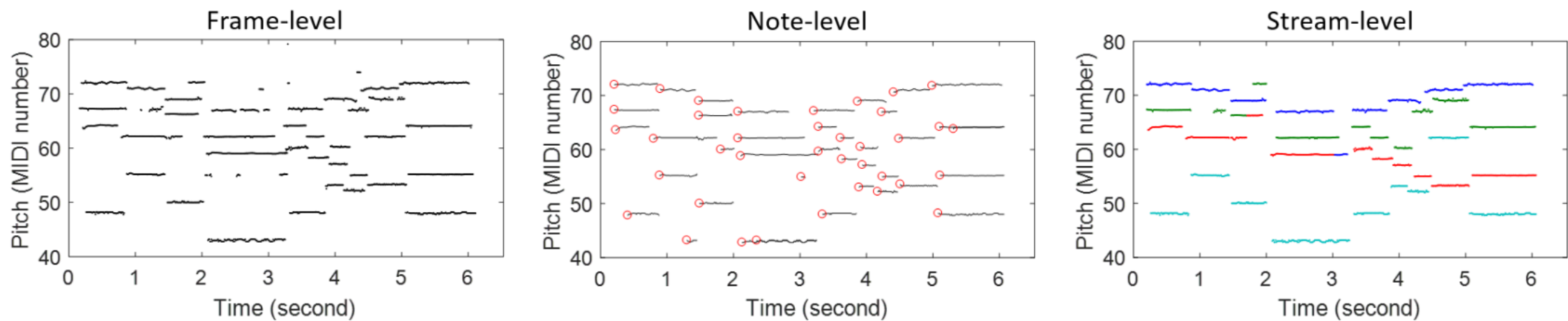
- **Deep Salience** (Bittner et al., 2017)
 - github.com/rabitt/ismir2017-deepsalience
- **Onset and Frames** (Hawthorne et al., 2018)
 - Piano only
 - github.com/jongwook/onsets-and-frames
- **Basic Pitch** (Bittner et al., 2022)
 - github.com/spotify/basic-pitch
 - basicpitch.spotify.com

Rachel M. Bittner, Brian McFee, Justin Salamon, Peter Li, and Juan P. Bello, "Deep Salience Representations for F0 Estimation in Polyphonic Music," *ISMIR*, 2017.

Curtis Hawthorne, Erich Elsen, Jialin Song, Adam Roberts, Ian Simon, Colin Raffel, Jesse Engel, Sageev Oore, and Douglas Eck, "Onsets and Frames: Dual-Objective Piano Transcription," *ISMIR*, 2018.

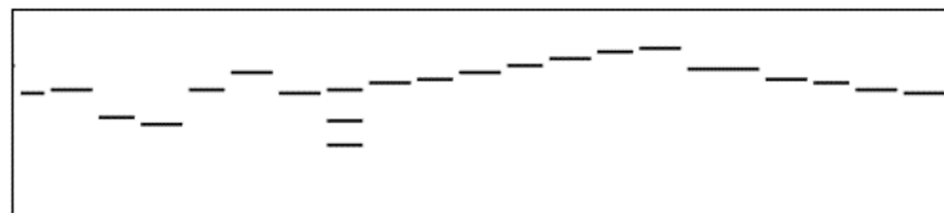
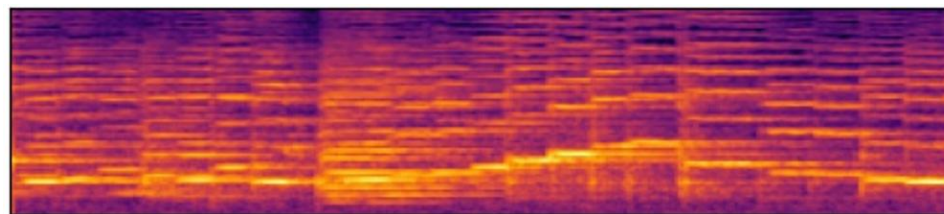
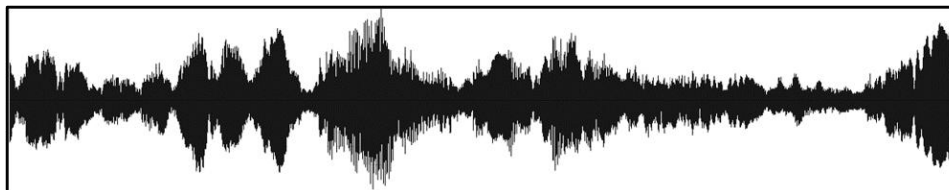
Rachel M. Bittner, Juan José Bosch, David Rubinstein, Gabriel Meseguer-Brocal, and Sebastian Ewert, "A Lightweight Instrument-Agnostic Model for Polyphonic Note Transcription and Multipitch Estimation," *ICASSP*, 2022.

F0 Estimation vs Music Transcription



(Source: Benetos et al., 2019)

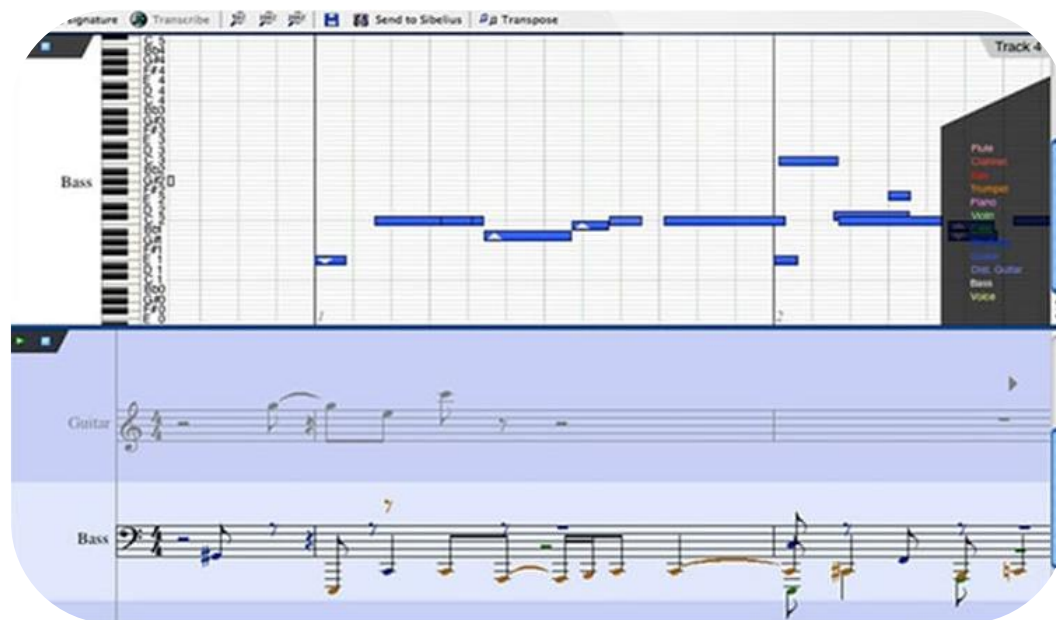
Music Transcription



(Source: Dong et al., 2022)

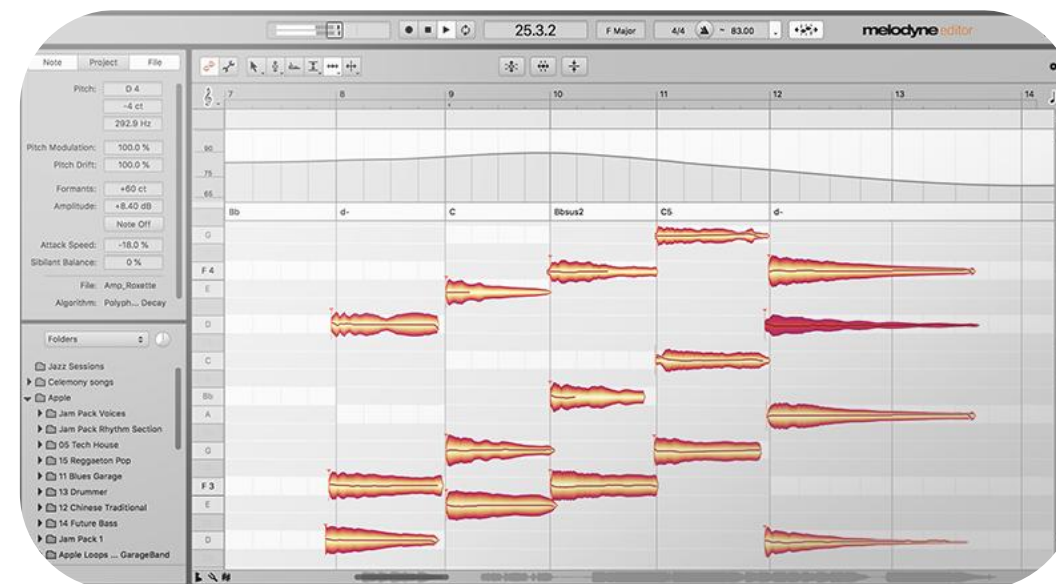
Commercial Music Transcription Software

AudioScore in Sibelius



(Source: Avid)

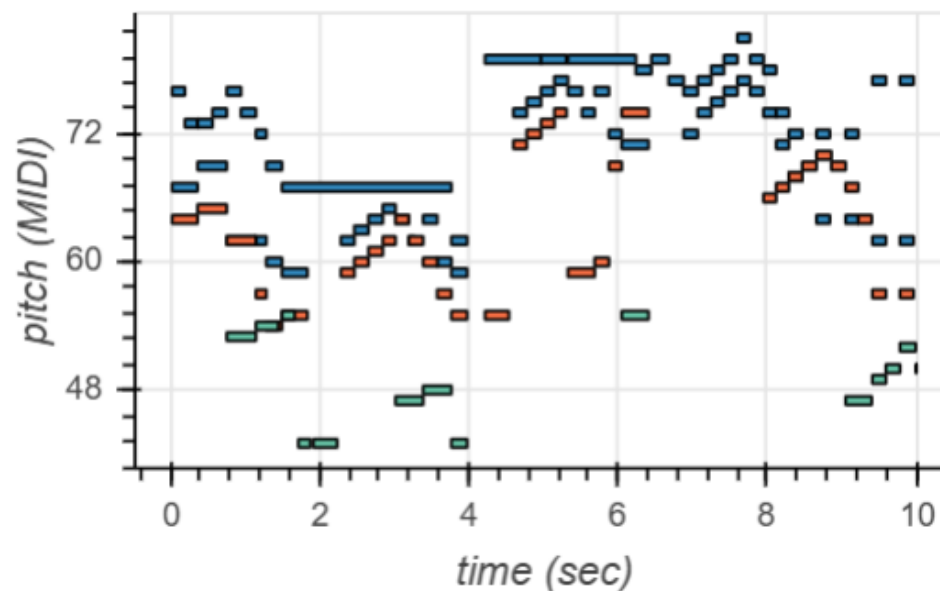
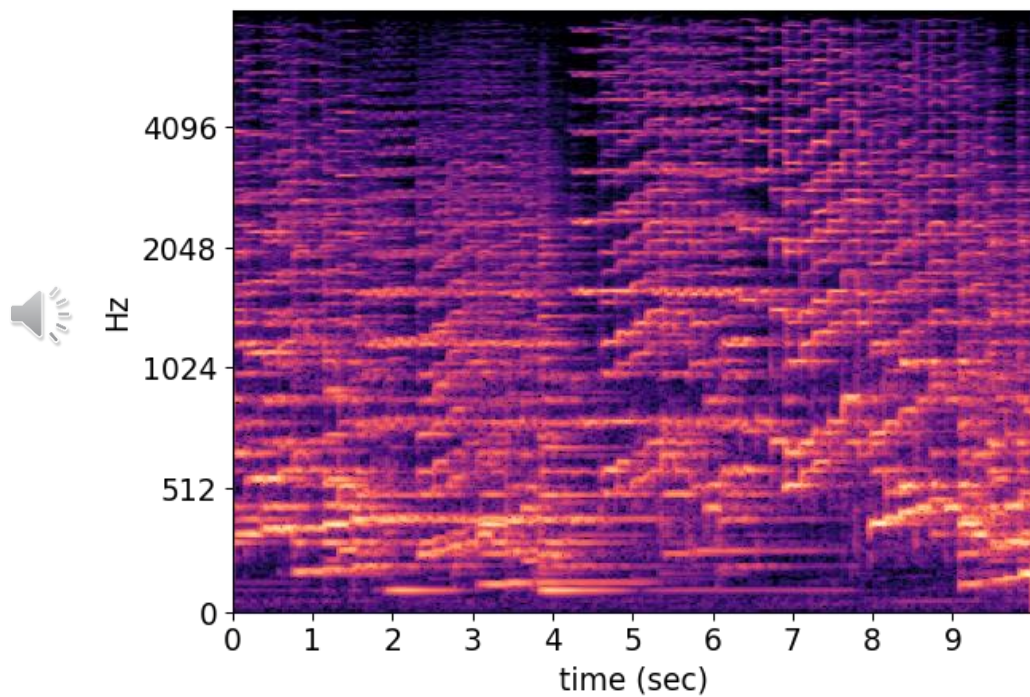
Melodyne Editor



(Source: Celemony)

Multitrack Transcription Models

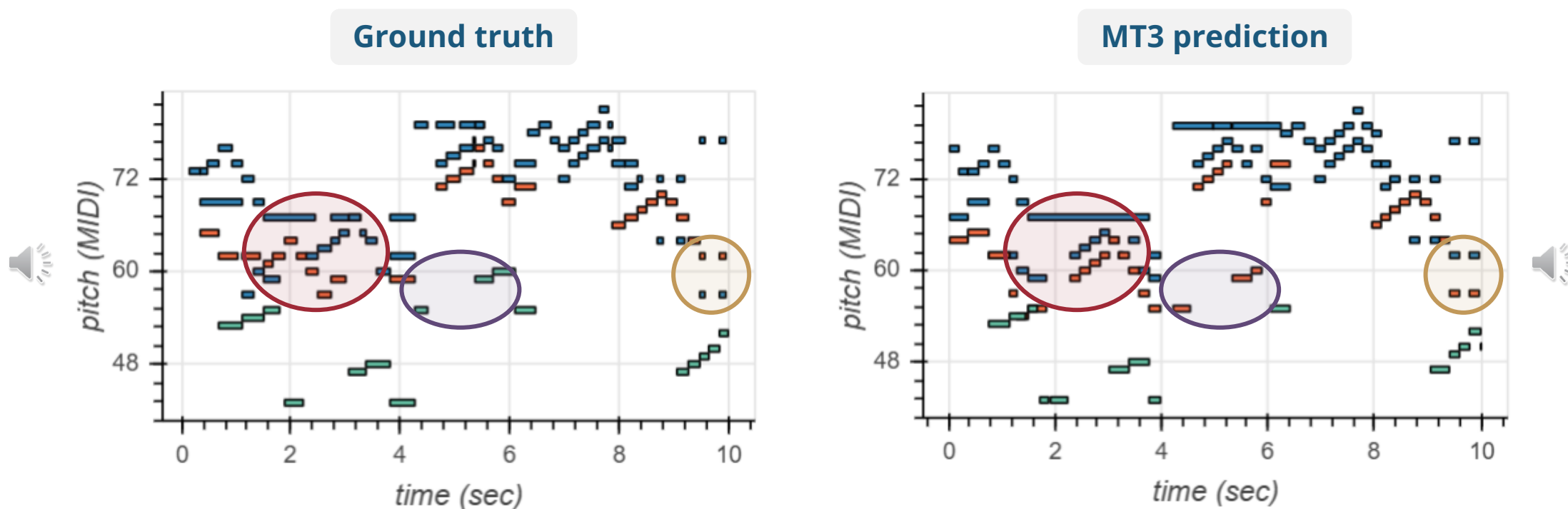
- **MT3** (Gardner et al., 2022)
 - github.com/magenta/mt3



(Source: Gardner et al., 2022)

Multitrack Transcription Models

- **MT3** (Gardner et al., 2022)
 - github.com/magenta/mt3



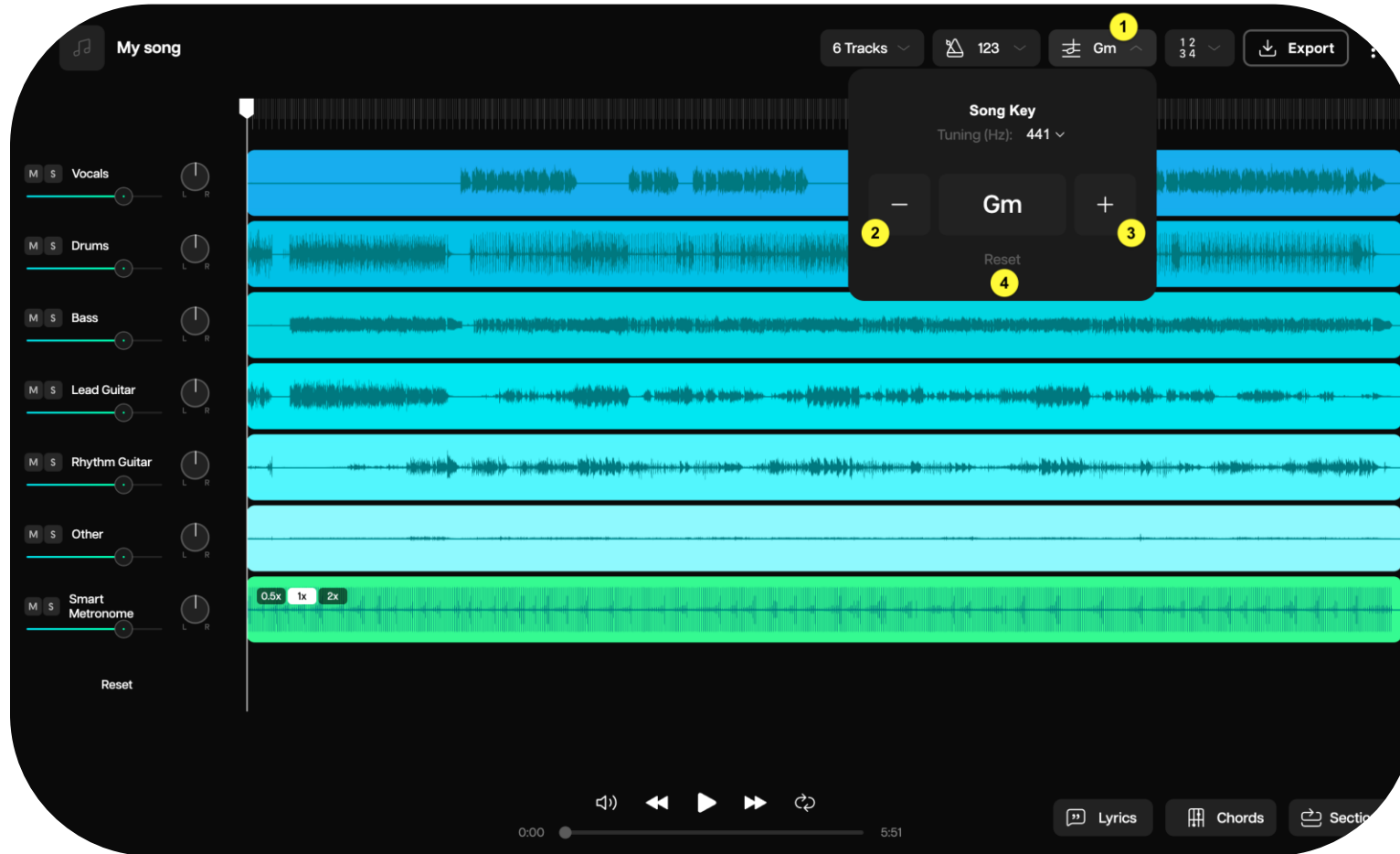
(Source: Gardner et al., 2022)

Resources

- Rachel Bittner, Mark Cartwright, and Ethan Manilow, “Programming MIR Baselines from Scratch: Three Case Studies,” *Tutorials of ISMIR*, 2021.
 - Part 1: [Transcription with NMF](#) (Ethan Manilow)
 - Part 2: [Pitch Tracking with pytorch](#) (Rachel Bittner)
 - Part 3: [Instrument Classification with OpenL3 & Tensorflow](#) (Mark Cartwright)
- Rachel Bittner, Alain de Cheveigné, and Johana Devaney, “Fundamental Frequency Estimation in Music,” *Tutorials of ISMIR*, 2018.
 - Part 1: [Pitch](#) (Alain de Cheveigné)
 - Part 2: [Polyphonic fundamental frequency estimation](#) (Rachel Bittner)
 - Part 3: [Applications](#) (Johana Devaney)

Harmony Analysis

Key Detection in Moises



(Source: Moises)

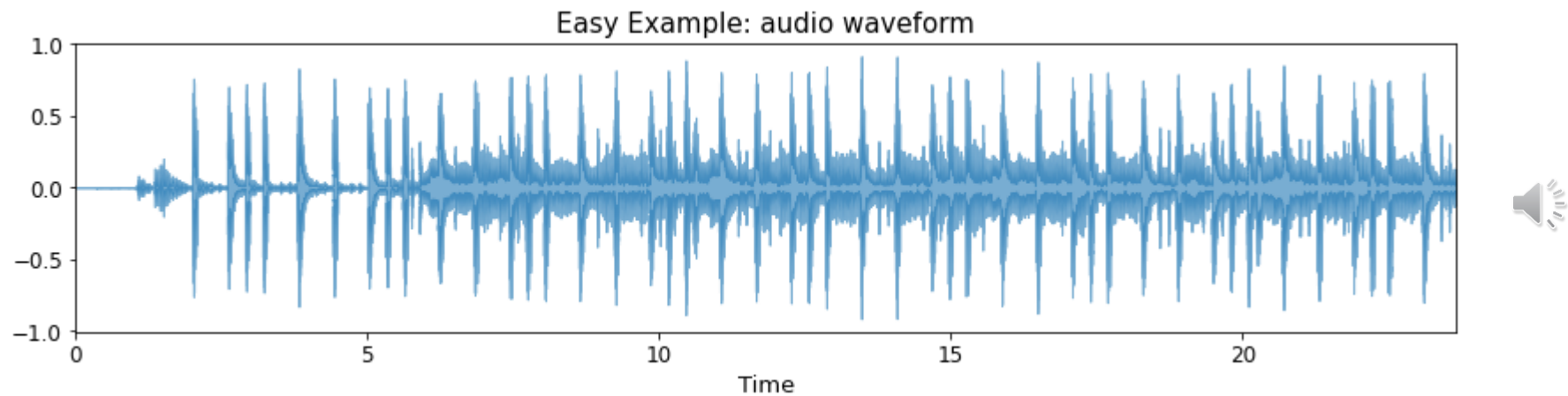
Chord Detection



youtube.com/shorts/_N3b_GARMfA

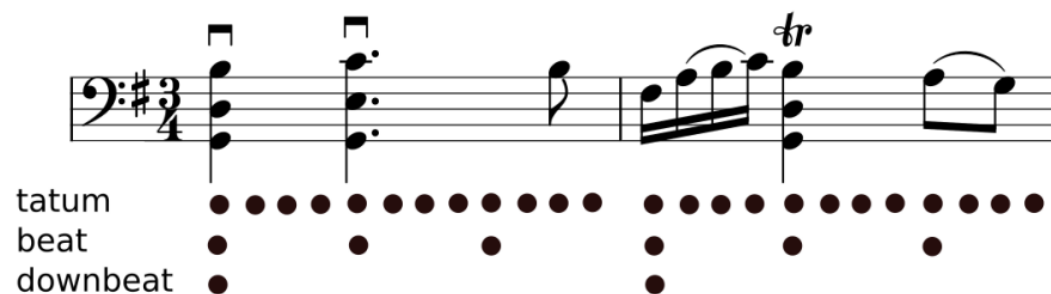
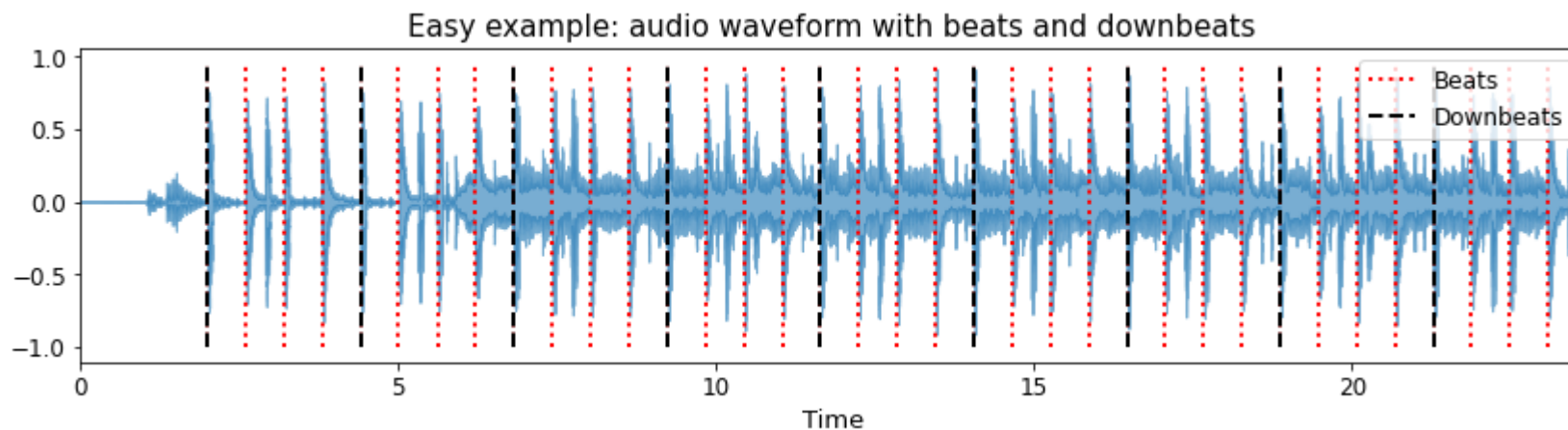
Rhythm Analysis

Beat & Downbeat Estimation



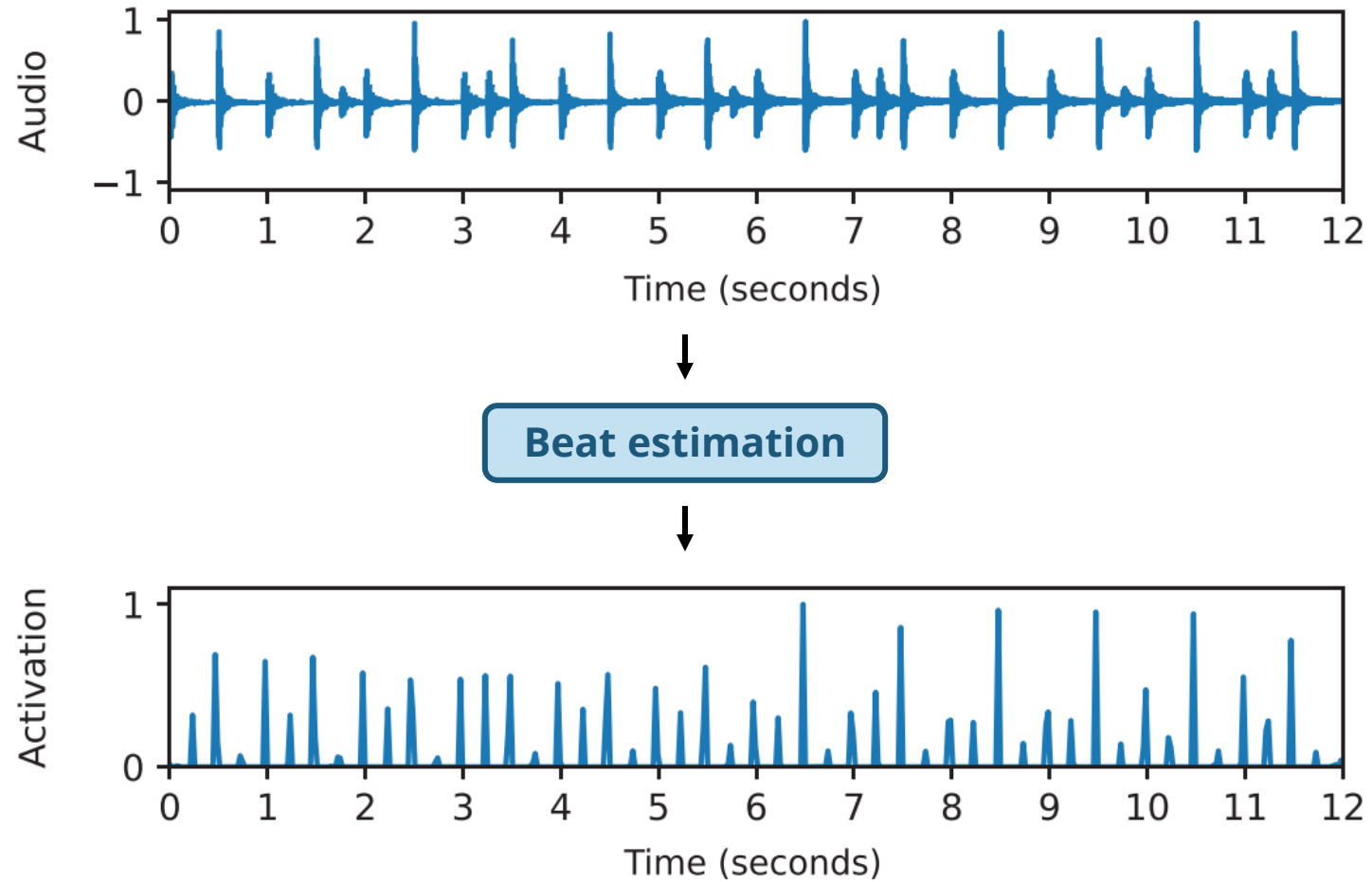
(Source: Davies et al., 2021)

Beat & Downbeat Estimation



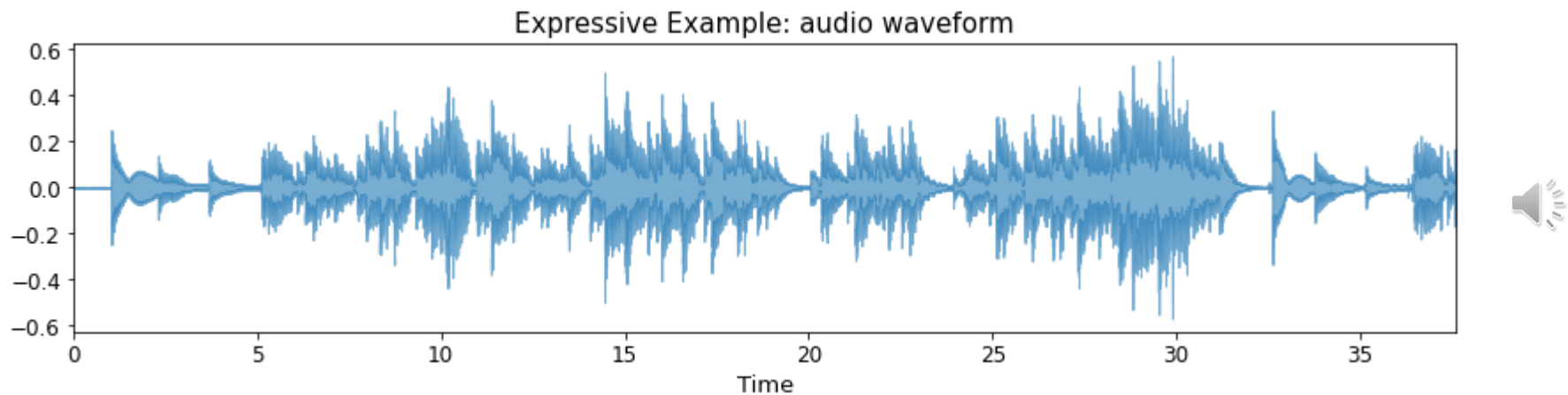
(Source: Davies et al., 2021)

Beat & Downbeat Estimation



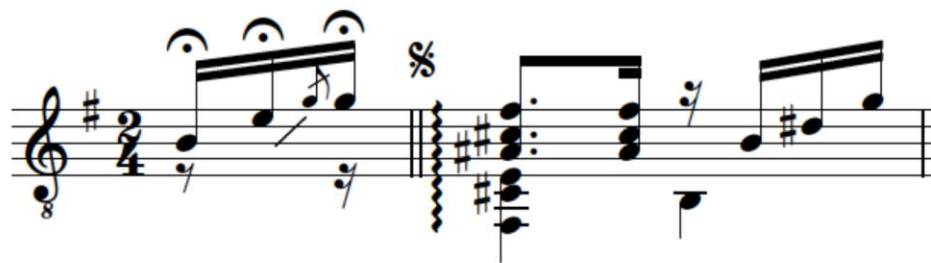
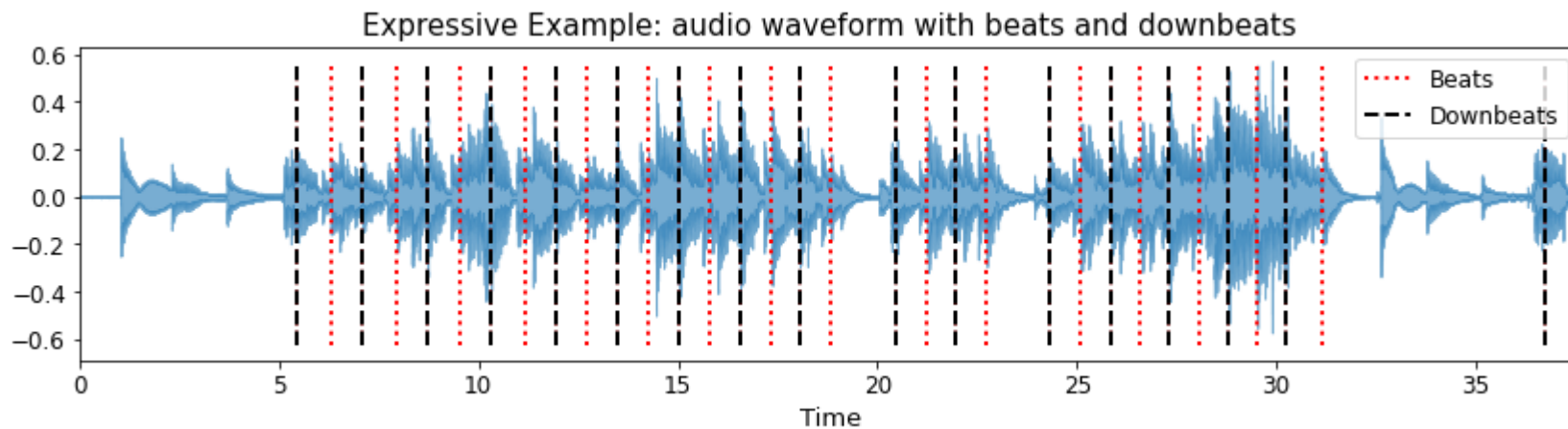
(Source: Meier et al., 2024)

Beat & Downbeat Estimation



(Source: Davies et al., 2021)

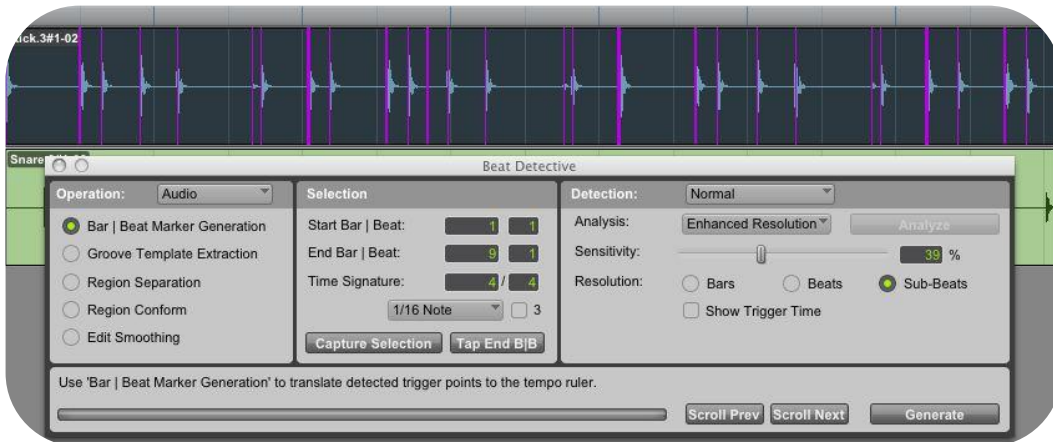
Beat & Downbeat Estimation



(Source: Davies et al., 2021)

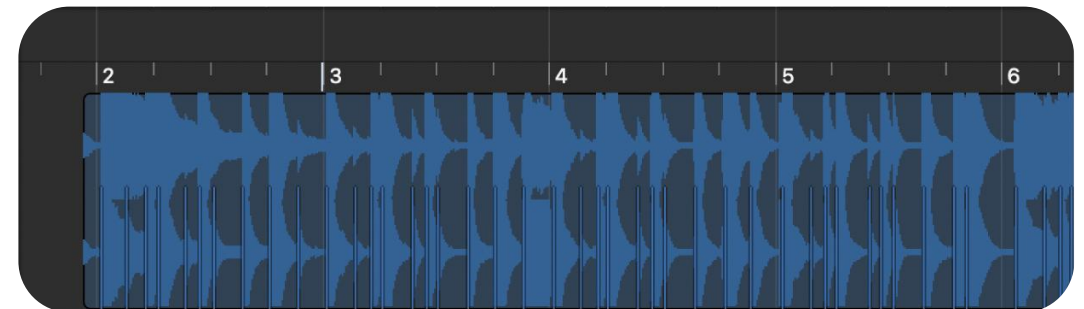
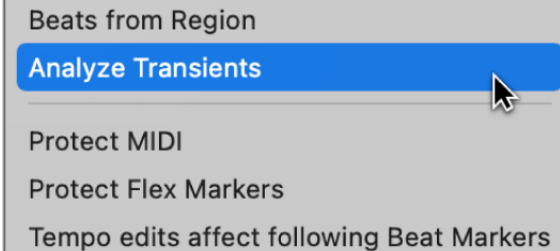
Beat Tracking in Pro Tools & Logic Pro

Beat Detective in Pro Tools



(Source: Logic Pro User Guide)

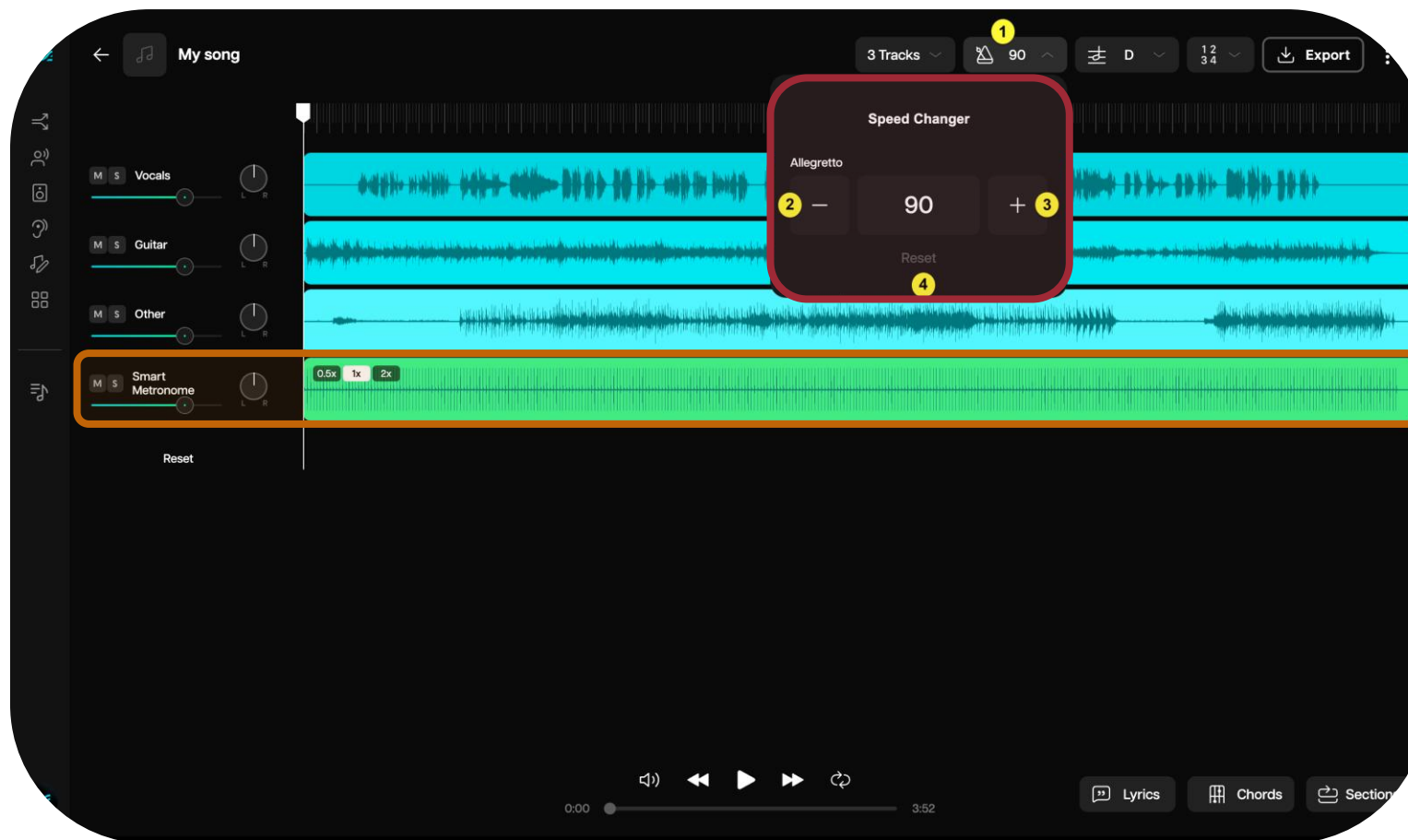
Beat Mapping in Logic Pro



(Source: Logic Pro User Guide)

Tempo Estimation & Beat Tracking in Moises

Tempo estimation



Beat tracking

(Source: Moises)

Resources on Rhythm Analysis

- Matthew E. P. Davies, Sebastian Böck, and Magdalena Fuentes, "[Tempo, Beat and Downbeat Estimation](#)," *Tutorials of ISMIR*, 2021.
- Hendrik Schreiber, Julián Urbano, and Meinard Müller, "[Music Tempo Estimation: Are We Done Yet?](#)," *TISMIR*, 3(1):111-125, 2020.

Structure Analysis

Structure Analysis

Music segmentation

The musical score is divided into sections labeled A1, A2, B1, B2, C, A3, B3, B4, and D. The score includes various musical notations such as dynamics (f, ff, p, sf), tempo markings (Allegro, Vivace, in tempo, poco rit.), and articulation (V). The sections are color-coded: A1 and A2 are pink, B1 and B2 are green, C is light blue, A3 is pink, B3 and B4 are green, and D is yellow.

Figure 4.5 following [Müller, FMP, Springer 2015]

(Source: Müller & Zalkow, 2019)

Hierarchical music segmentation

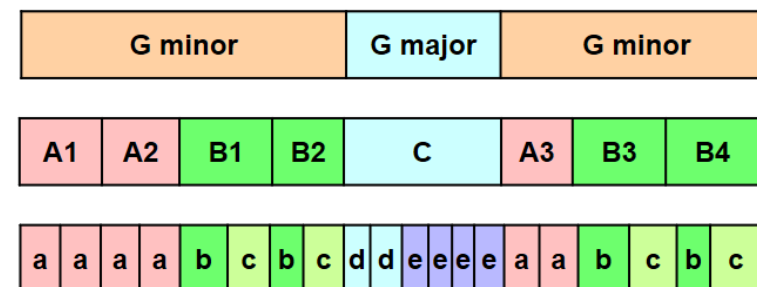
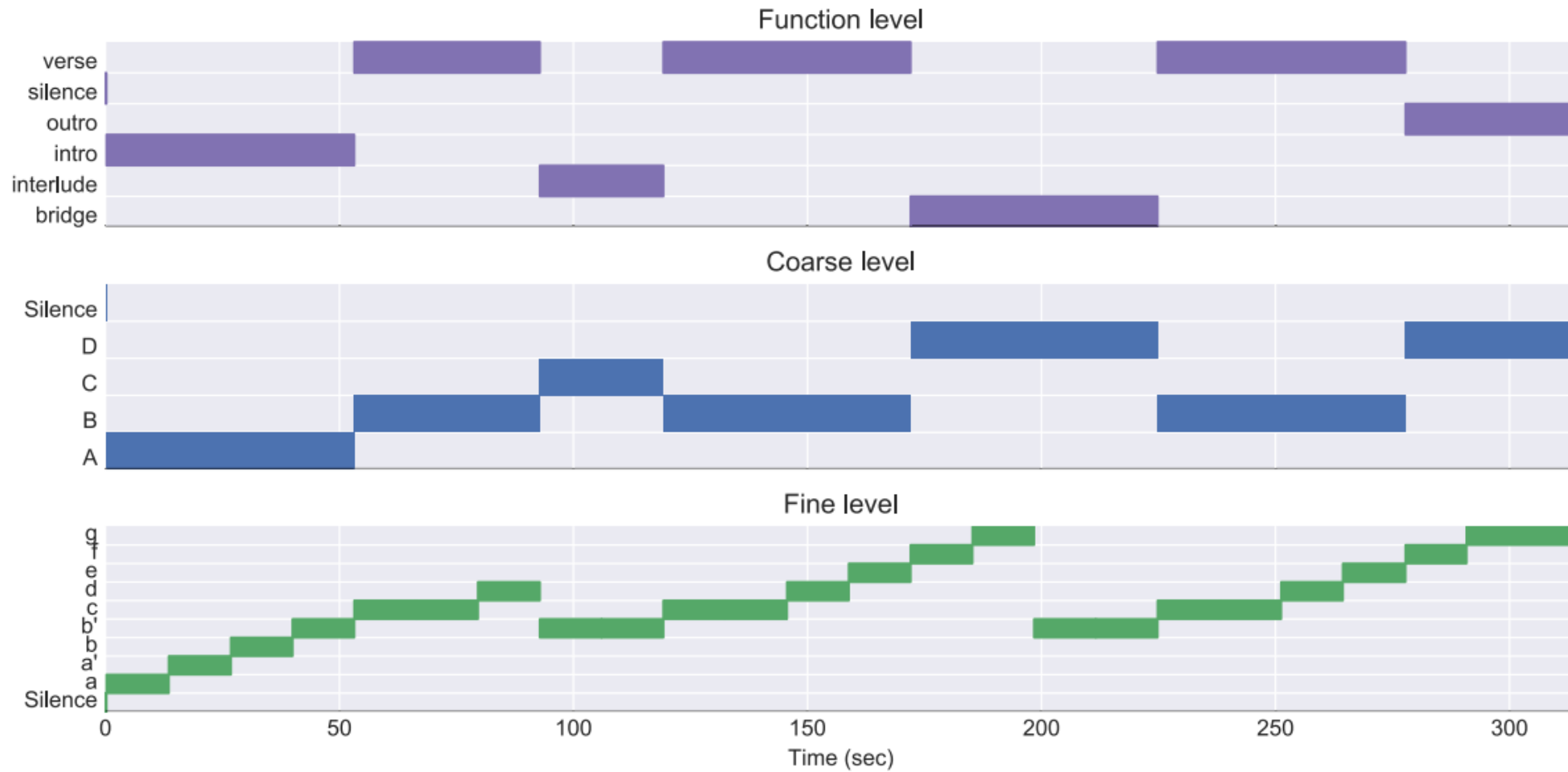


Figure 4.28 from [Müller, FMP, Springer 2015]

(Source: Müller & Zalkow, 2019)

Hierarchical Music Segmentation



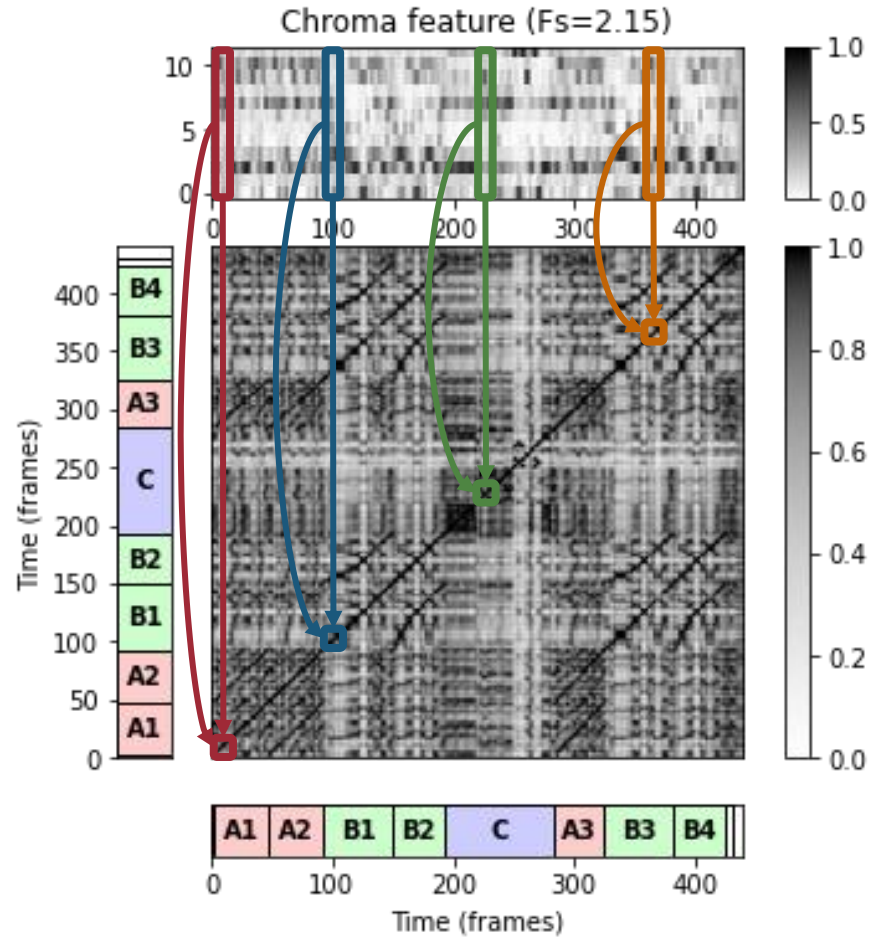
(Source: Müller & Zalkow, 2019)

1



(Source: Müller & Zalkow, 2019)

1



(Source: Müller & Zalkow, 2019)

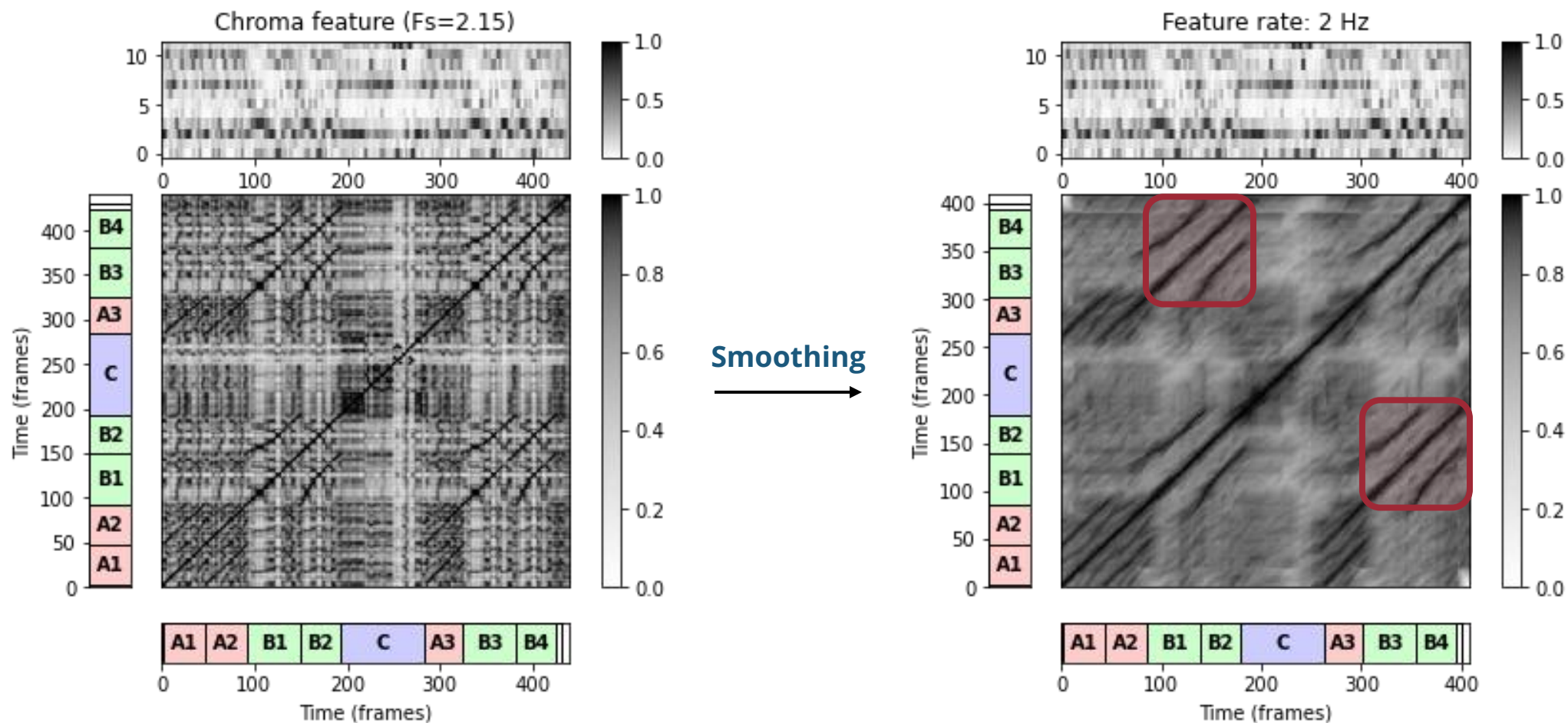


(Source: Müller & Zalkow, 2019)



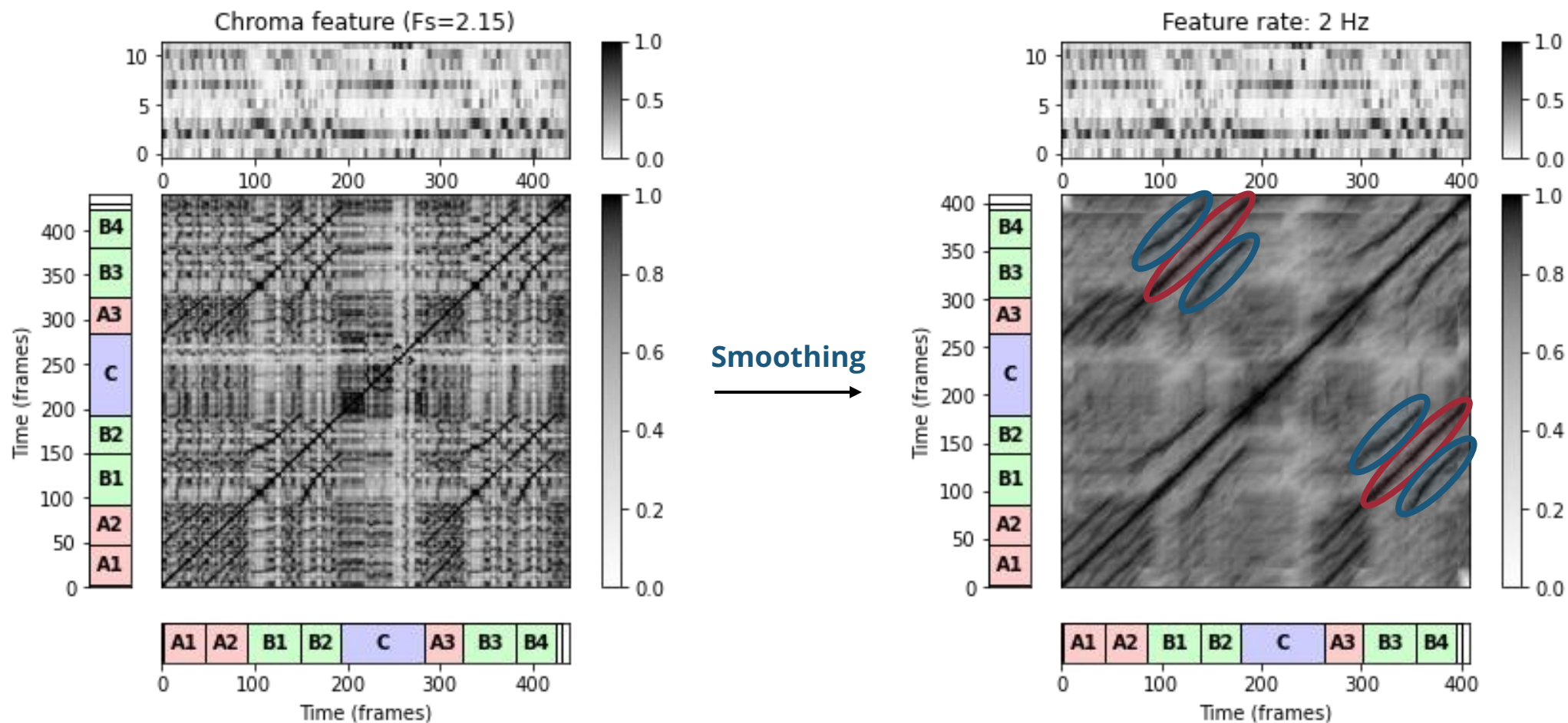
(Source: Müller & Zalkow, 2019)

Self-Similarity Matrices (SSMs)



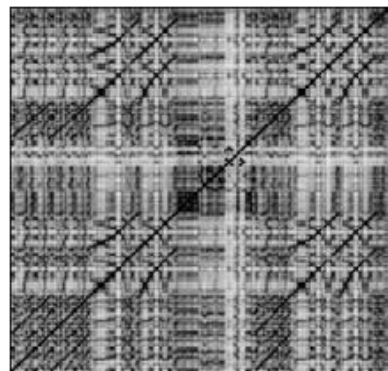
(Source: Müller & Zalkow, 2019)

Self-Similarity Matrices (SSMs)

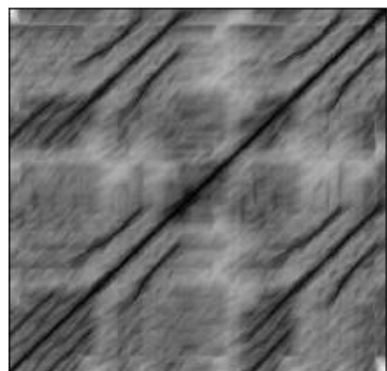


(Source: Müller & Zalkow, 2019)

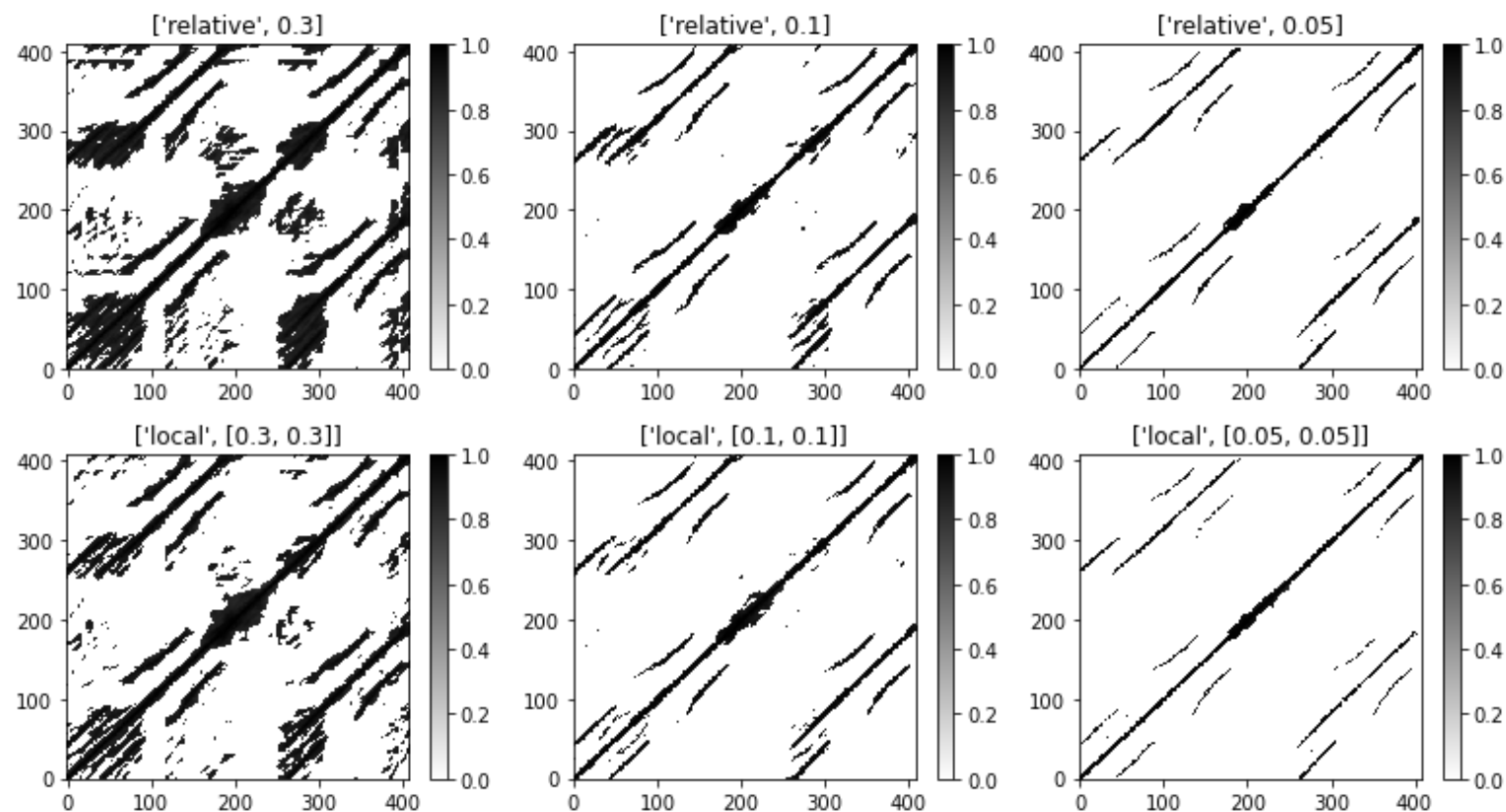
Self-Similarity Matrices (SSMs)



↓ Smoothing

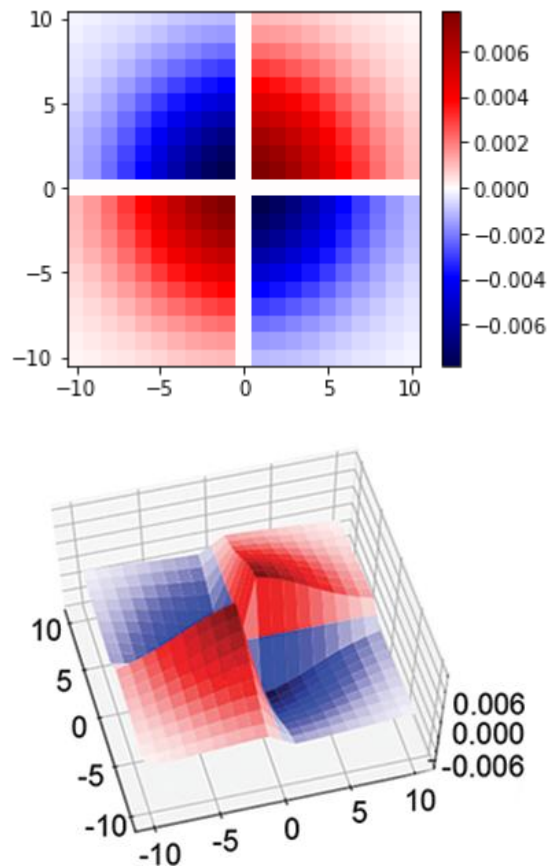


↗ Thresholding



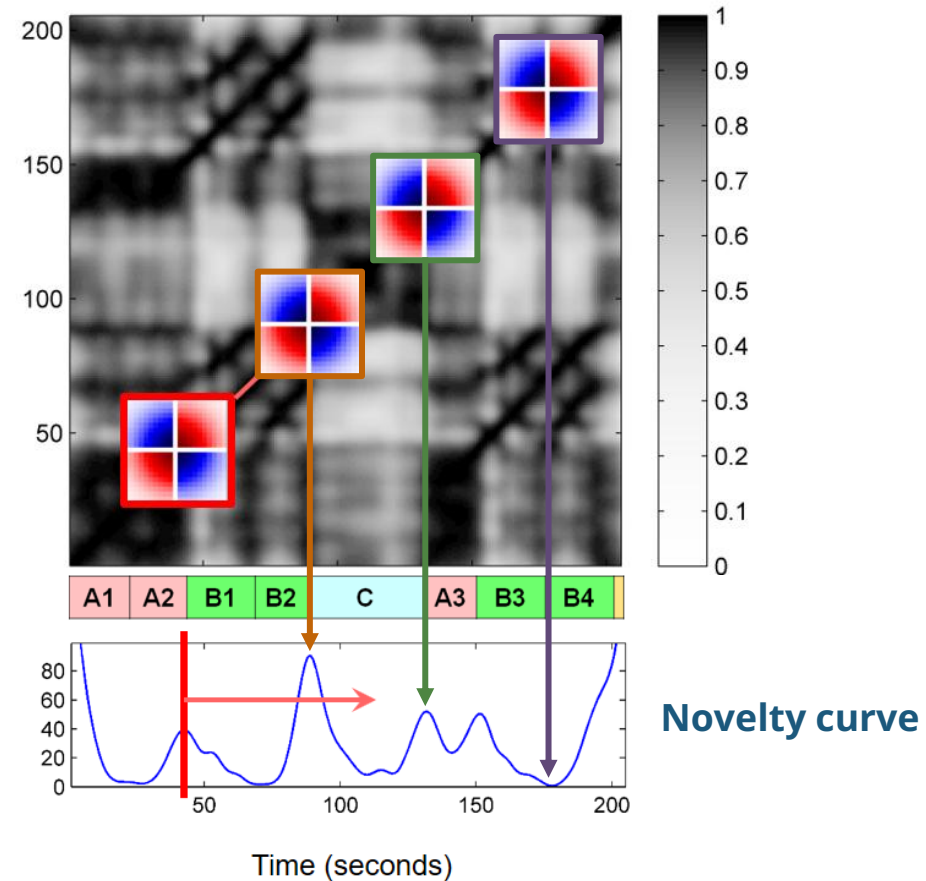
(Source: Müller & Zalkow, 2019)

Self-Similarity Matrices (SSMs)



(Source: Müller & Chiu, 2024)

Figure 4.24 from [Müller, FMP, Springer 2015]



(Source: Müller & Zalkow, 2019)

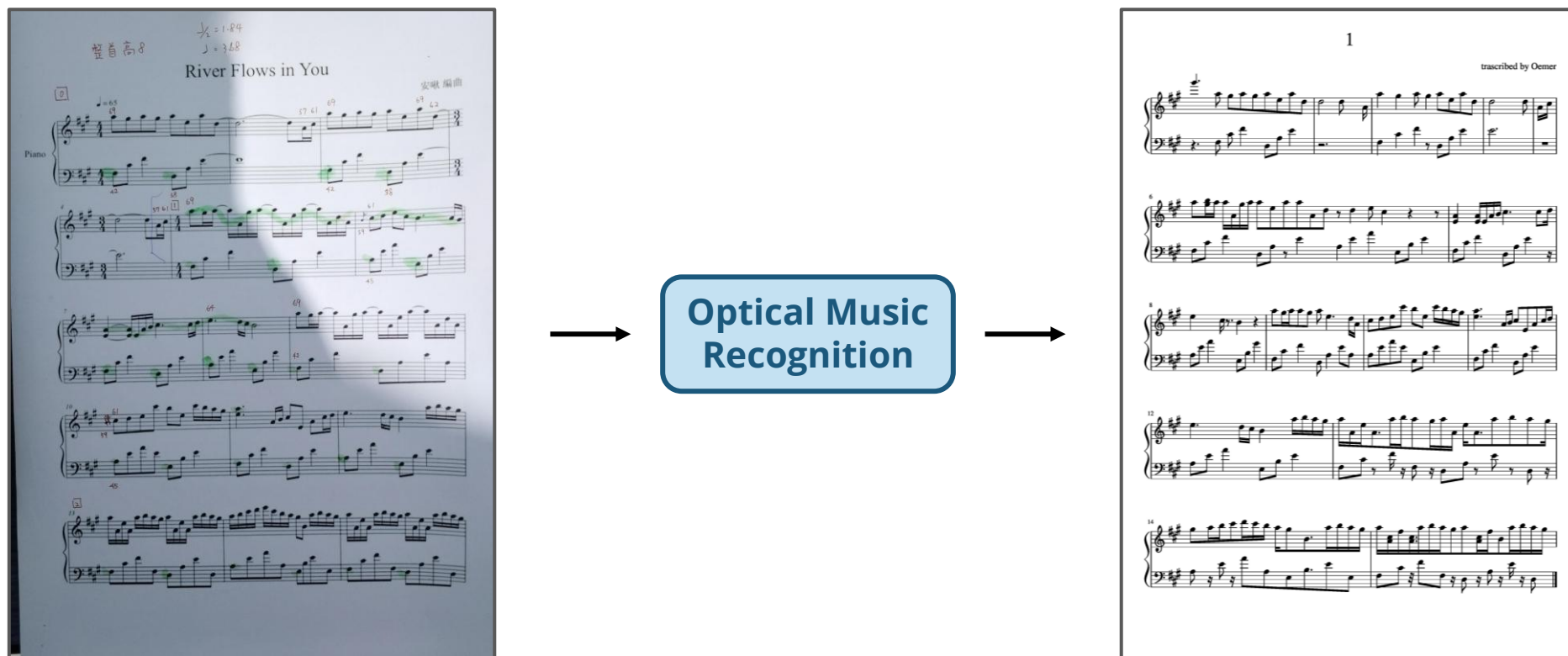
Resources on Music Structure Analysis

- Meinard Müller and Ching-Yu Chiu, "[A Basic Tutorial on Novelty and Activation Functions for Music Signal Processing](#)," *TISMIR*, 7(1):179-194, 2024.
- Oriol Nieto, Gautham J. Mysore, Cheng-i Wang, Jordan B. L. Smith, Jan Schlüter, Thomas Grill, and Brian McFee, "[Audio-Based Music Structure Analysis: Current Trends, Open Challenges, and Applications](#)," *TISMIR*, 3(1):246-263, 2020.
- Meinard Müller & Jordan B. L. Smith, "Music Structure Analysis," *Tutorials of ISMIR*, 2014. ([part 1](#), [part 2](#), [part 3](#))

Optical Music Recognition

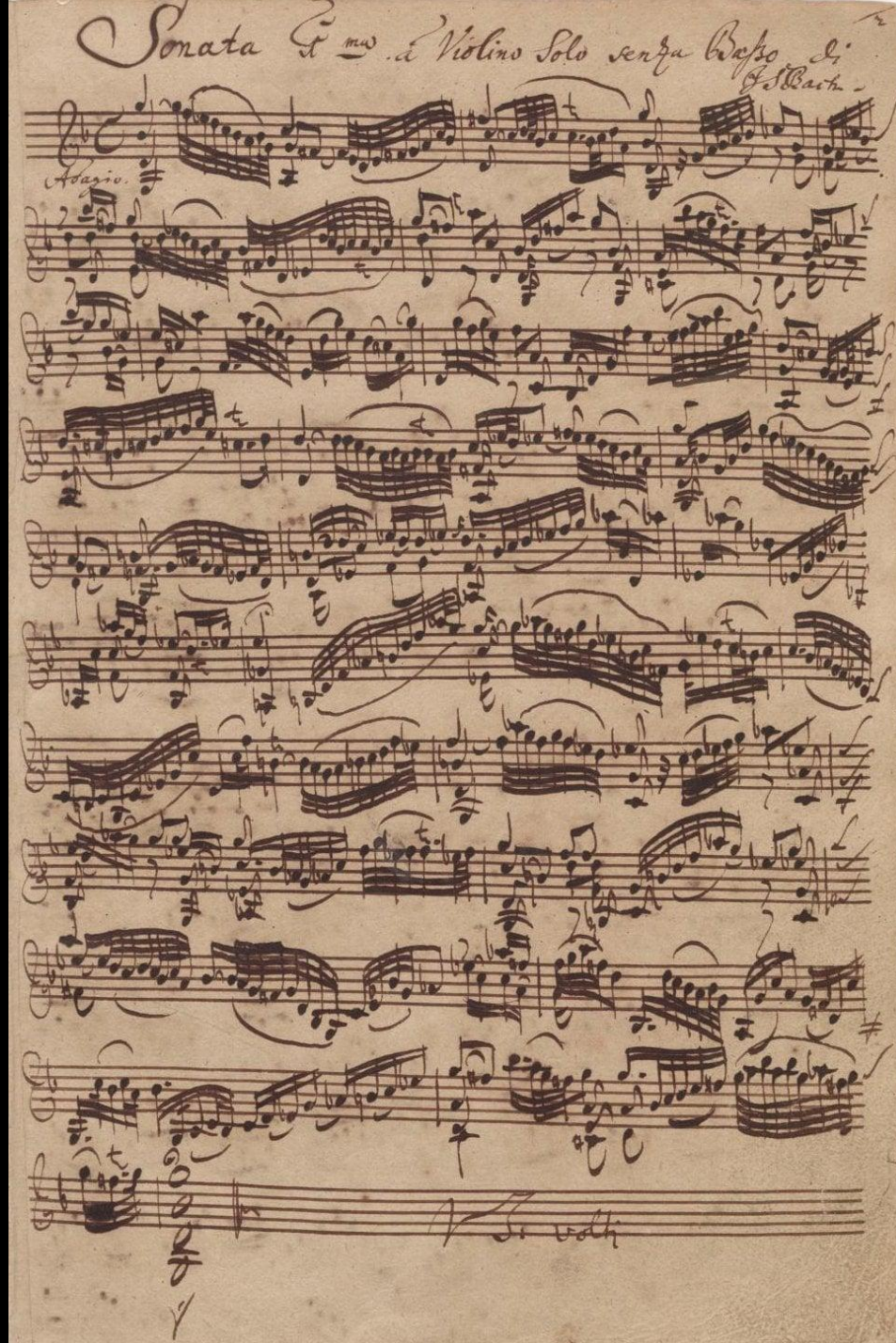
Optical Music Recognition (OMR)

- **Goal:** Convert **scanned sheet music** into **digital musical notation**



Challenges of OMR

Violin Sonata No. 1 in G minor
(BWV 1001)



| Challenges of OMR

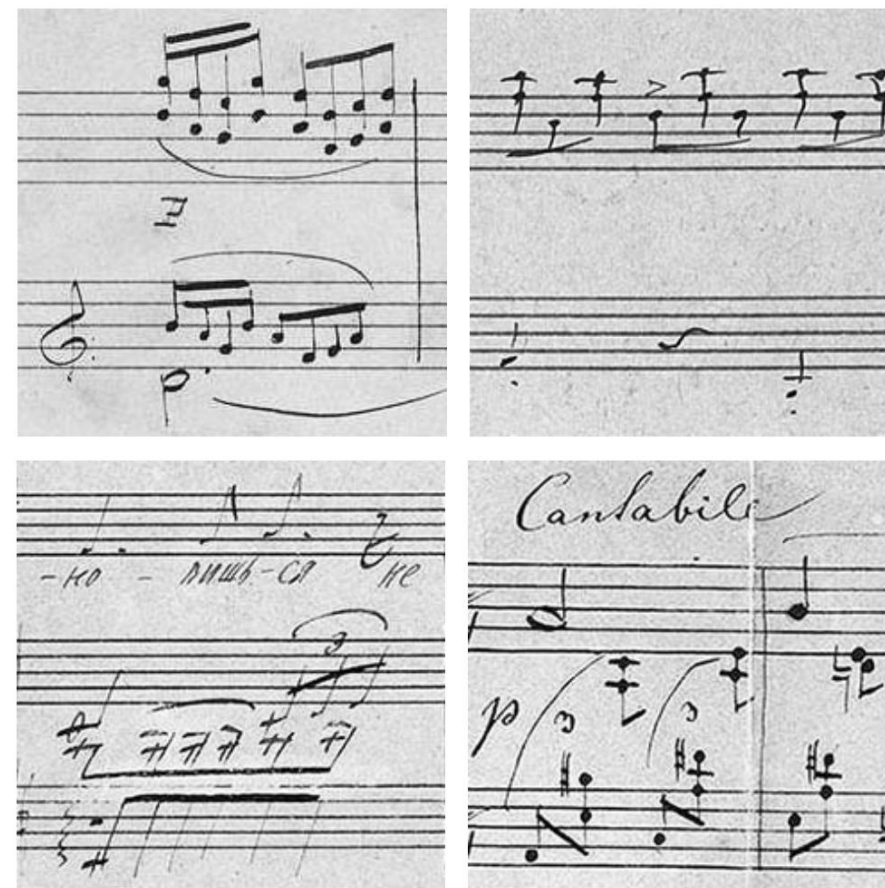
Weihnachtsoratorium
(Christmas Oratorio; BWV 248)



Challenges of OMR



(Source: Calvo-Zaragoza et al., 2018)

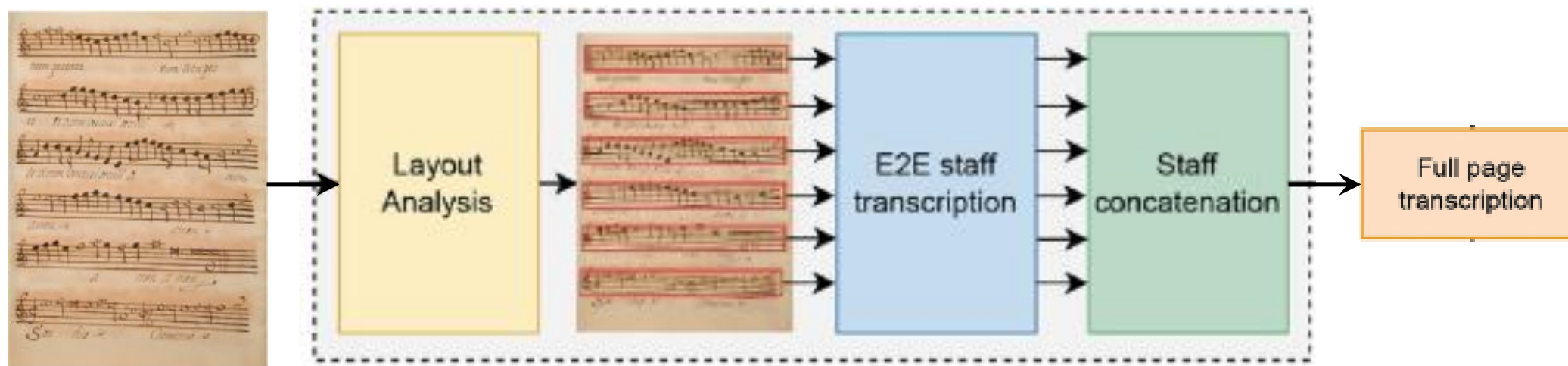


(Source: Novotný & Pokorný, 2015)

Jorge Calvo-Zaragoza, Juan C. Martinez-Sevilla, Carlos Penarrubia, and Antonio Rios-Vila, "Optical Music Recognition: Recent Advances, Current Challenges, and Future Directions," *ICDAR*, 2023.

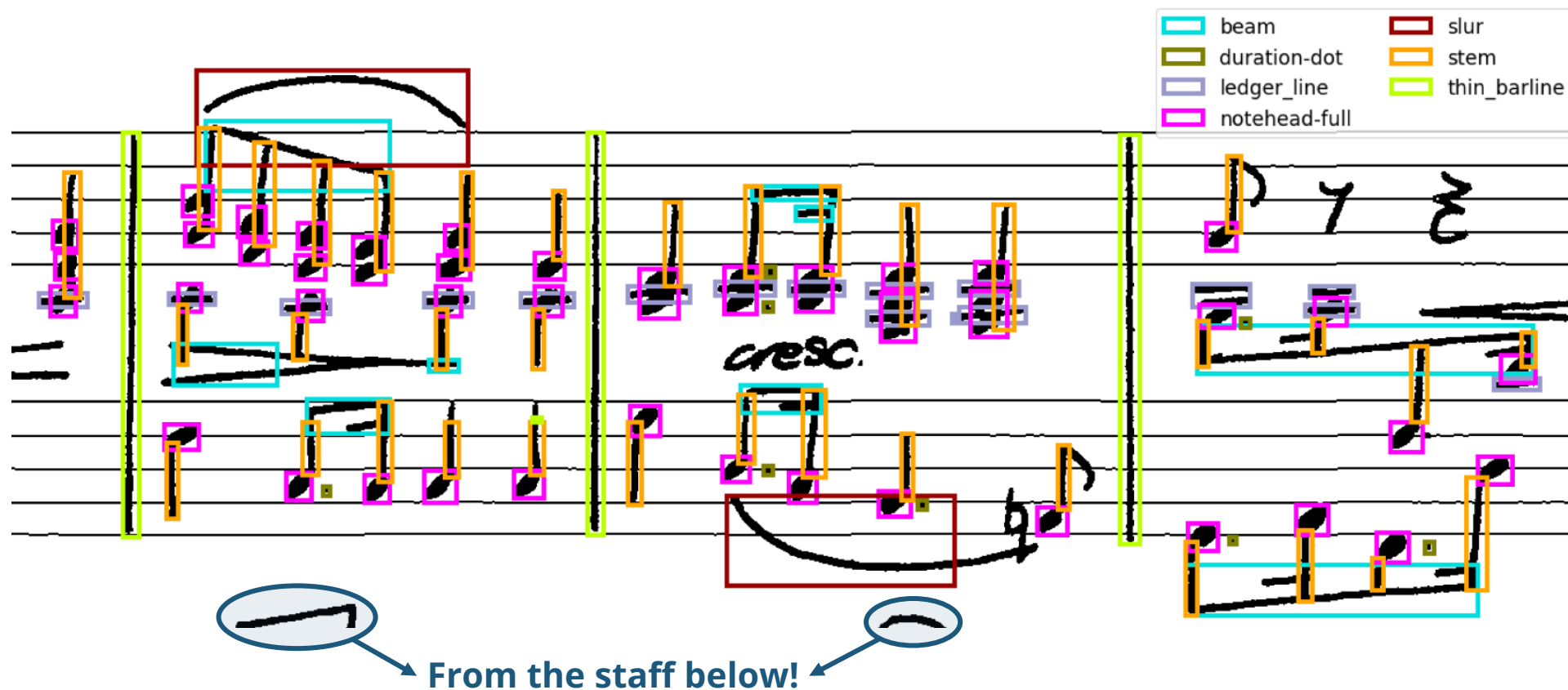
Jiří Novotný and Jaroslav Pokorný, "Introduction to Optical Music Recognition: Overview and Practical Challenges," *DATESO*, 2015.

Common Pipeline of OMR Systems



(Source: Calvo-Zaragoza et al., 2018)

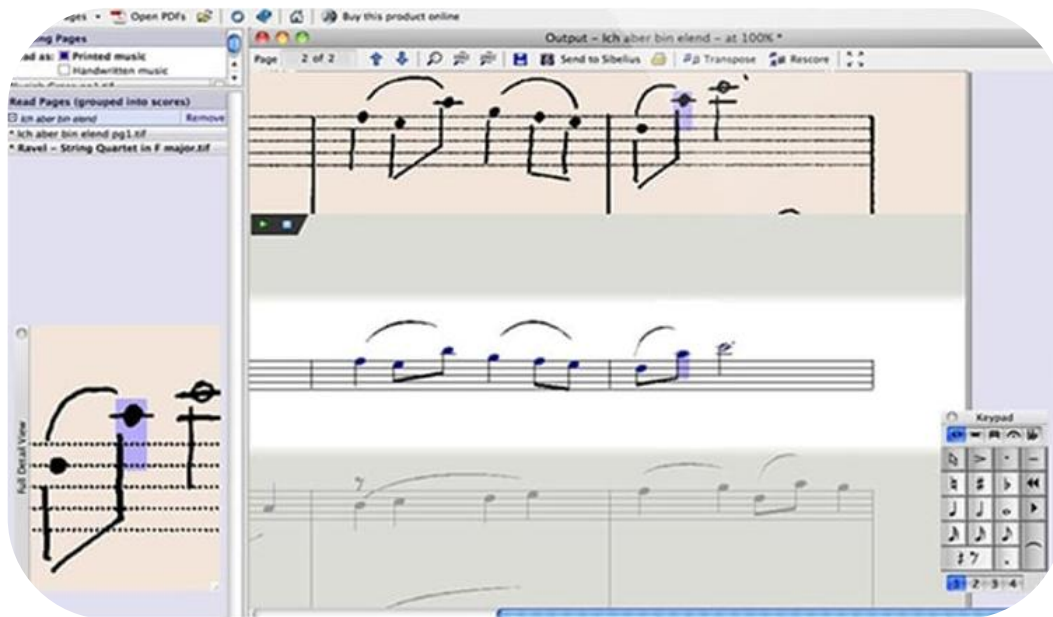
Musical Object Recognition



(Source: Pacha et al., 2018)

Commercial OMR Software

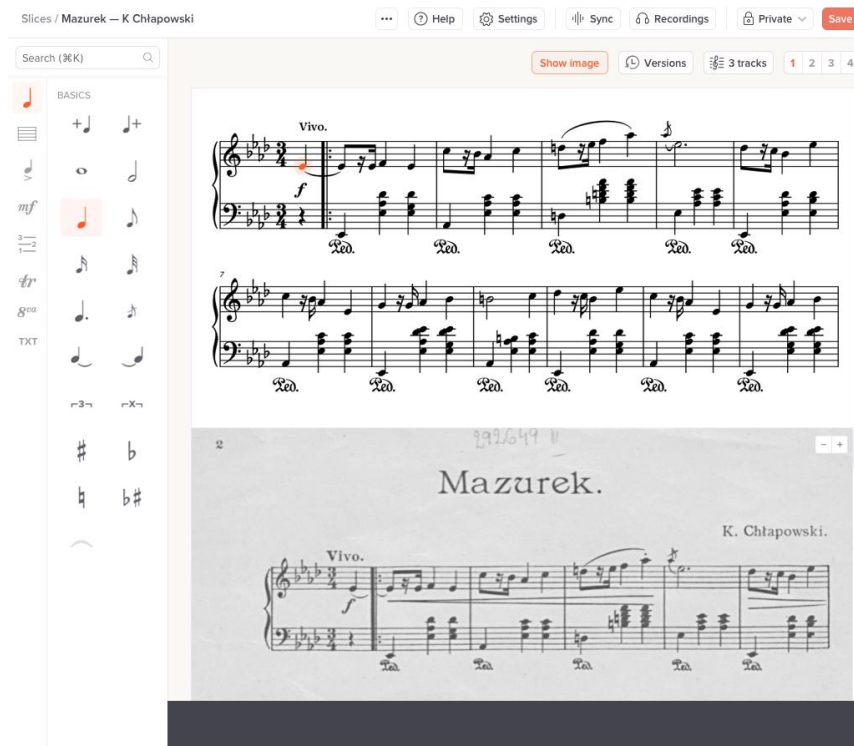
PhotoScore & NotateMe in Sibelius



(Source: Avid)

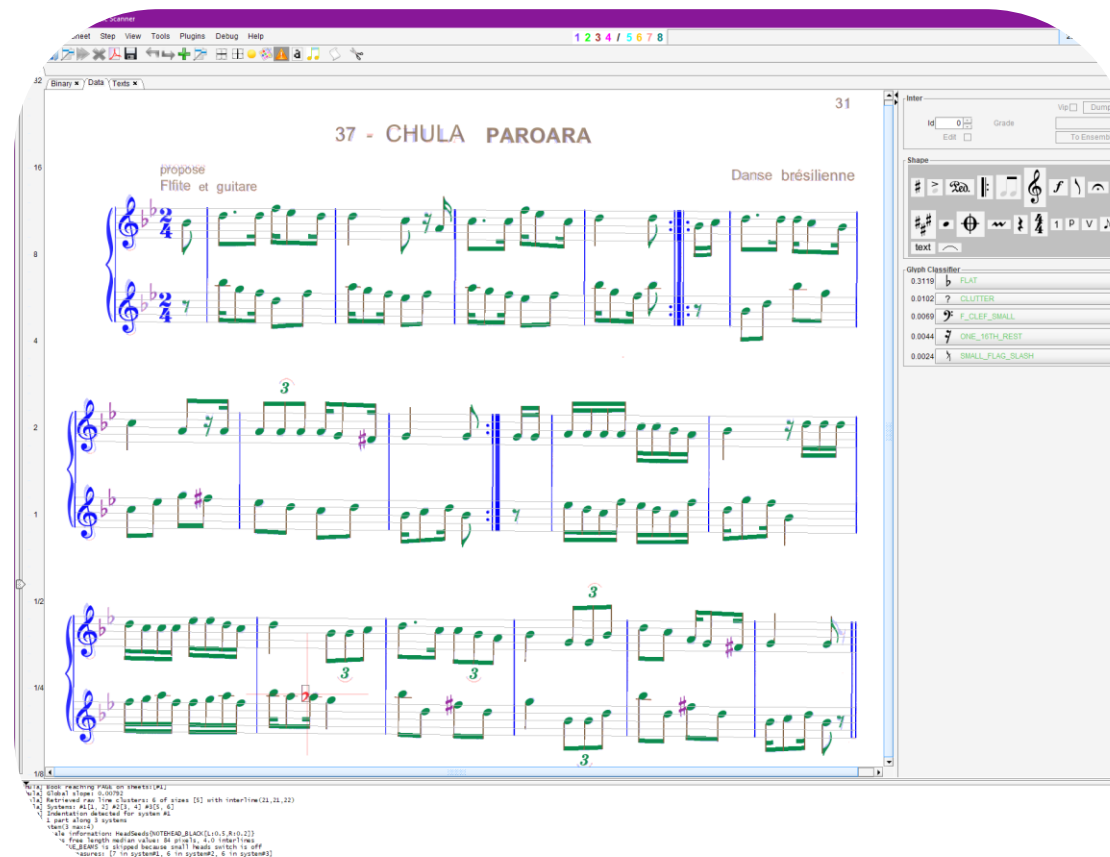
avid.com/products/photoscore-and-notateme-ultimate
soundslice.com/sheet-music-scanner/

Soundslice



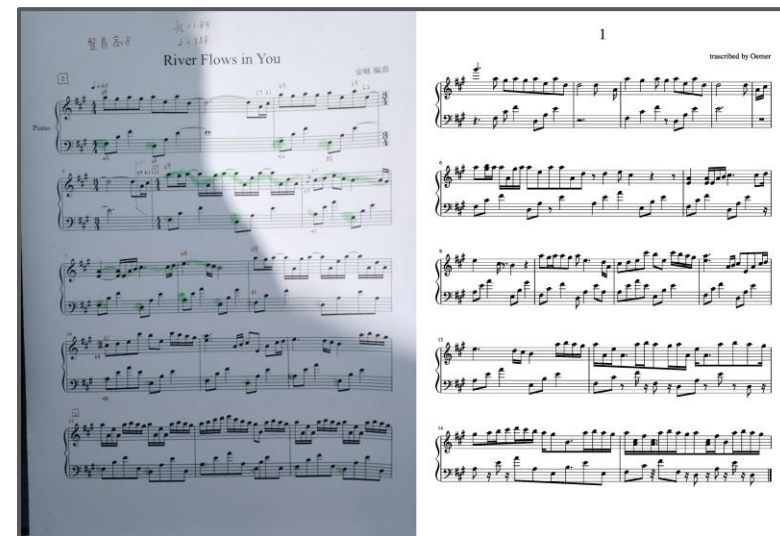
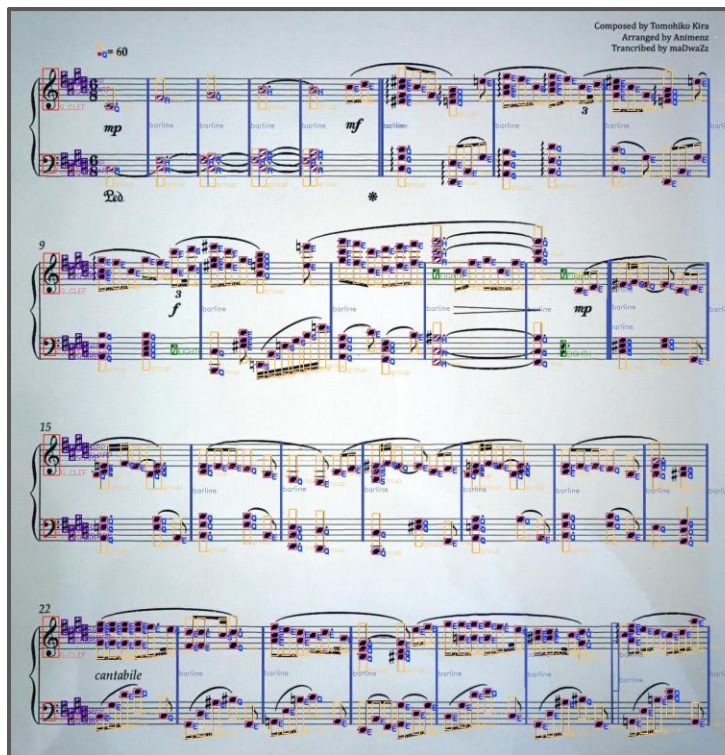
(Source: Soundslice)

Open-source OMR Software: Audiveris



github.com/Audiveris/audiveris

Open-source OMR Software: Oemer



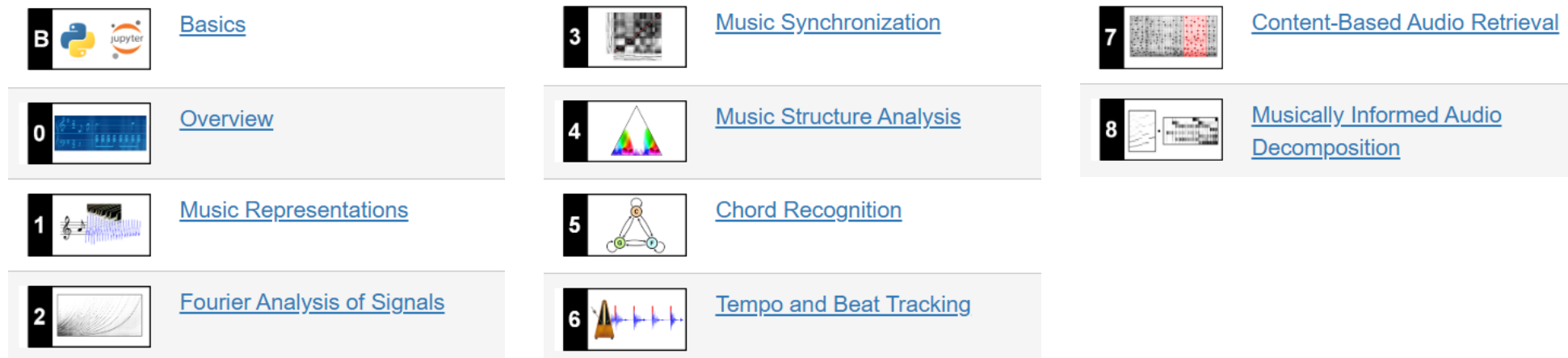
github.com/BreezeWhite/oemer

Resources on Optical Music Recognition (OMR)

- Jorge Calvo-Zaragoza, Jan Hajič jr., Alexander Pacha, and Ichiro Fujinaga, "[Optical Music Recognition for Dummies](#)," *Tutorials of ISMIR*, 2021. ([slides](#))
- OMR Datasets: apacha.github.io/OMR-Datasets/

Resources on Music Information Research (MIR)

- Meinard Müller, “[Fundamentals of Music Processing – Using Python and Jupyter Notebooks](#),” *Springer*, 2021.
- Meinard Müller and Frank Zalkow, “[FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing](#),” *ISMIR*, 2019.
 - Jupyter notebooks available at audiolabs-erlangen.de/FMP



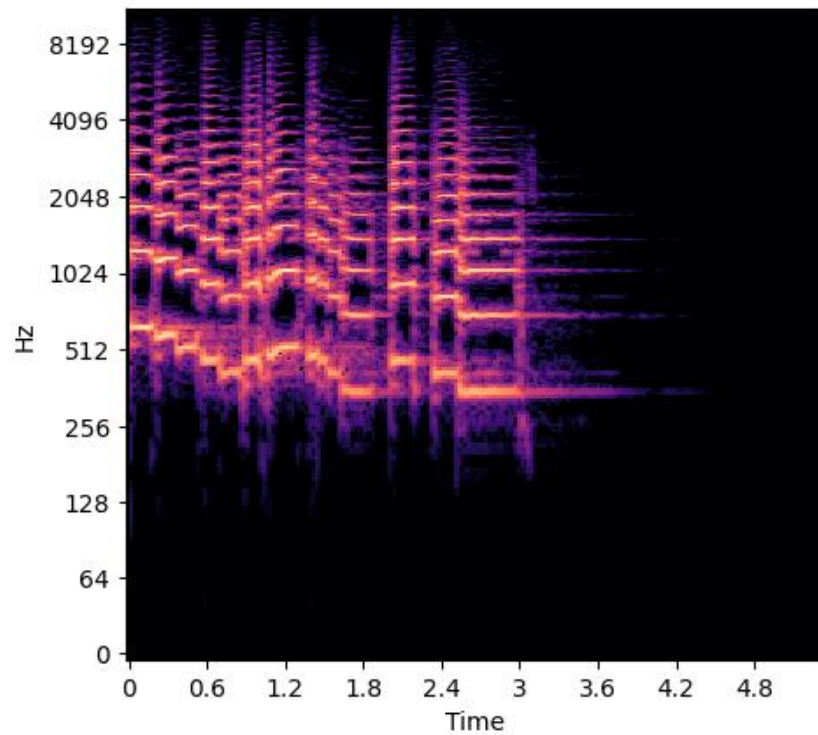
(Source: Müller & Zalkow, 2019)

Resources on Music Information Research (MIR)

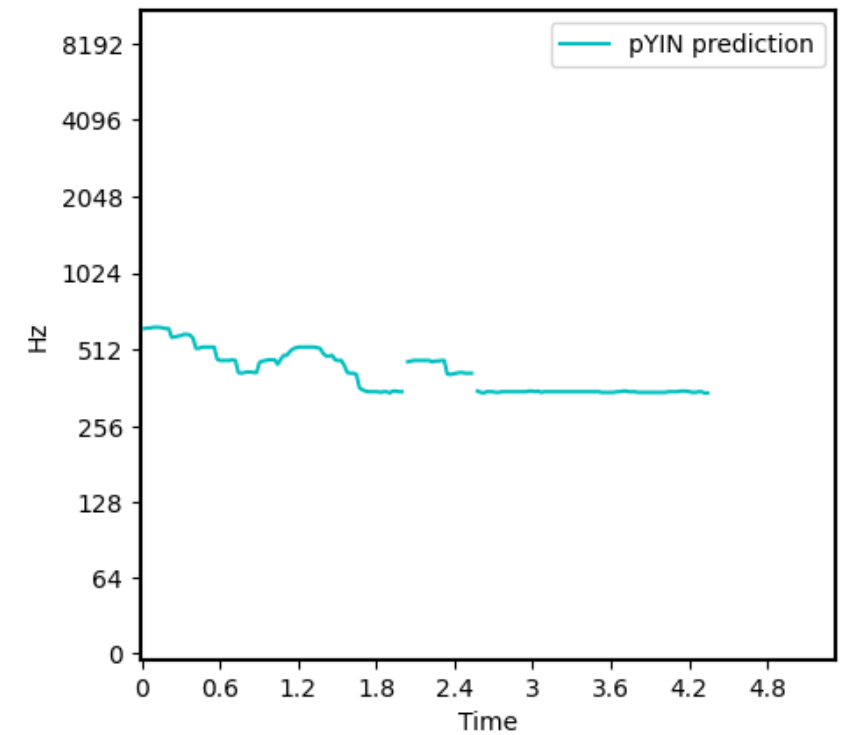
- Masataka Goto, Jin Ha Lee, and Meinard Müller, "[Exploring 25 Years of Music Information Retrieval: Perspectives and Insights](#)," *Tutorials of ISMIR*, 2024.
- Geoffroy Peeters, Gabriel Meseguer-Brocal, Alain Riou, and Stefan Lattner, "[Deep Learning 101 for Audio-based MIR](#)," *Tutorials of ISMIR*, 2024. ([book](#))
- Keunwoo Choi, György Fazekas, Kyunghyun Cho, and Mark Sandler, "[A Tutorial on Deep Learning for Music Information Retrieval](#)," *arXiv preprint arXiv:1709:04396*, 2017. ([code](#))

Recap

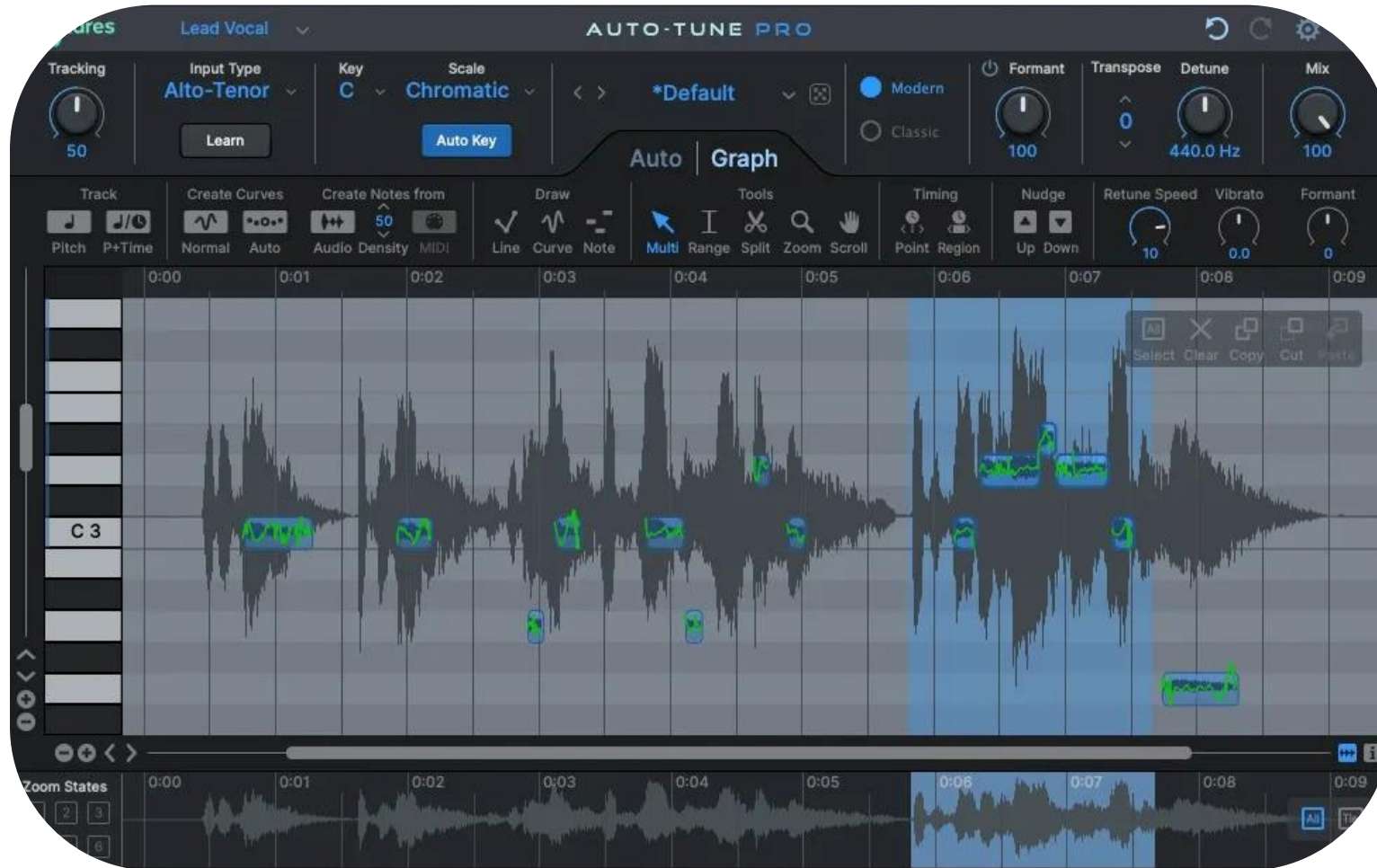
Fundamental Frequency (F0) Estimation



→ **F0 Estimation** →

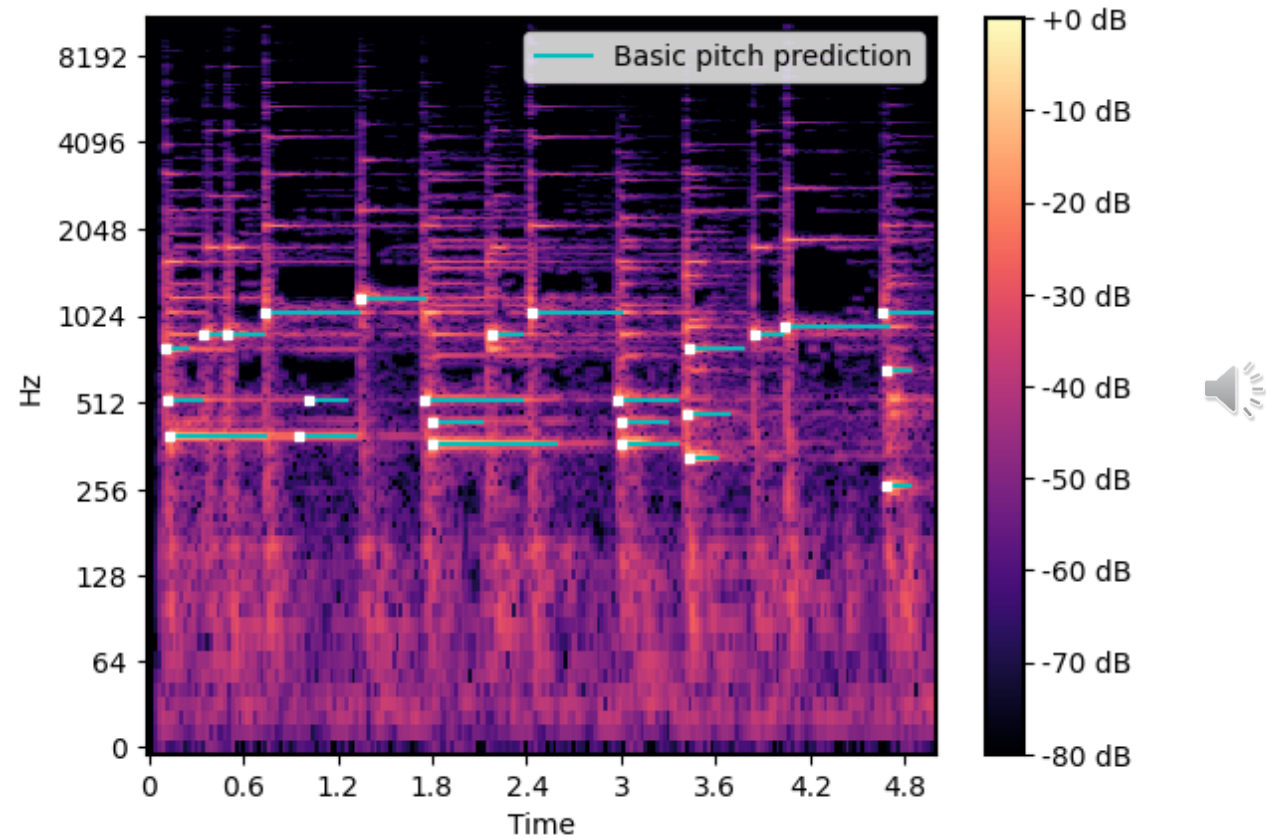


Auto-tune Pro



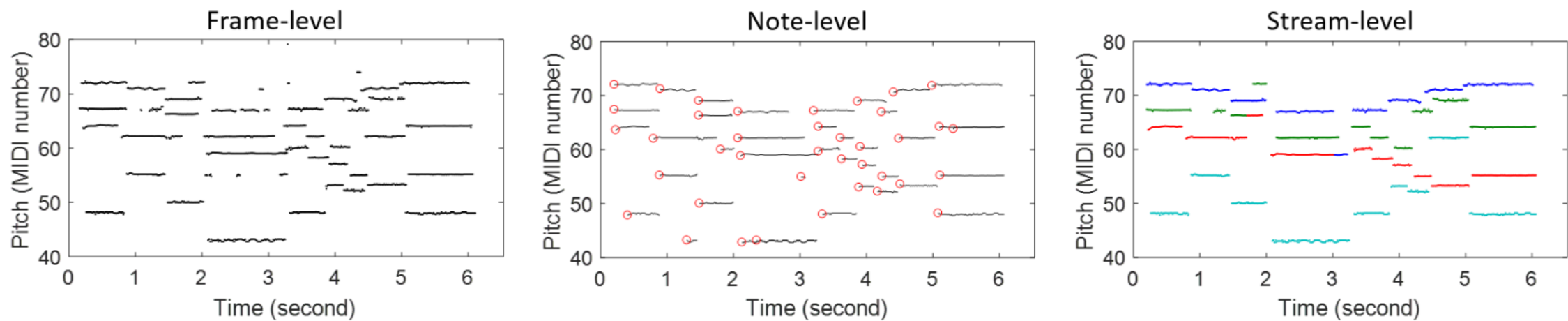
(Source: Antares Audio Technologies)

Polyphonic F0 Estimation



basicpitch.spotify.com

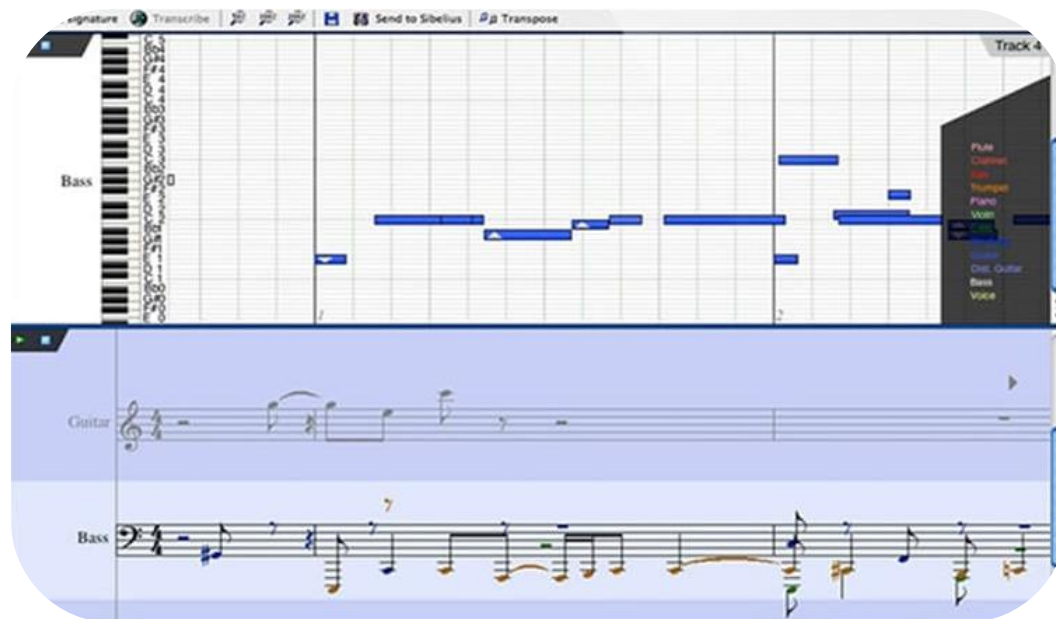
F0 Estimation vs Music Transcription



(Source: Benetos et al., 2019)

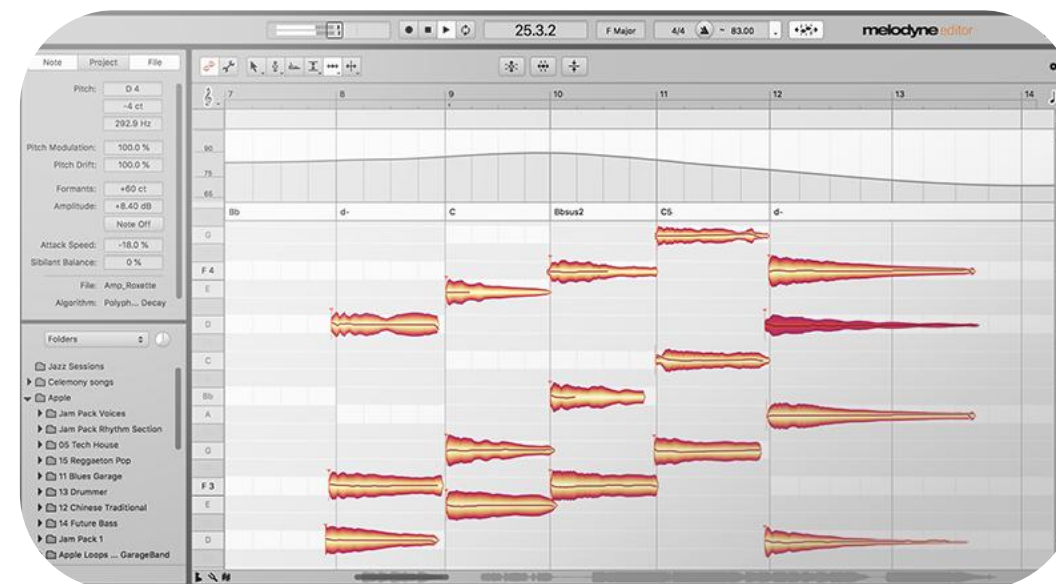
Commercial Music Transcription Software

AudioScore in Sibelius



(Source: Avid)

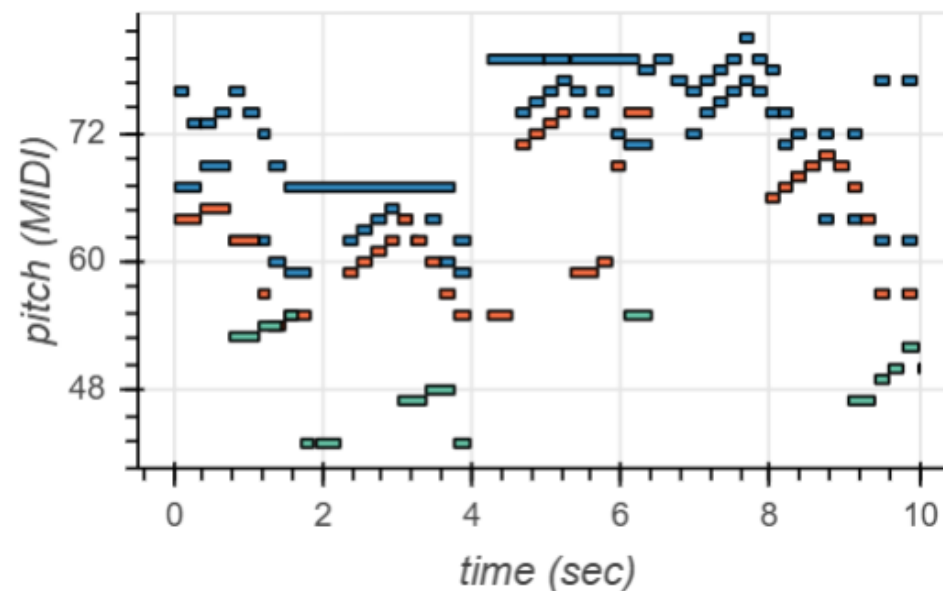
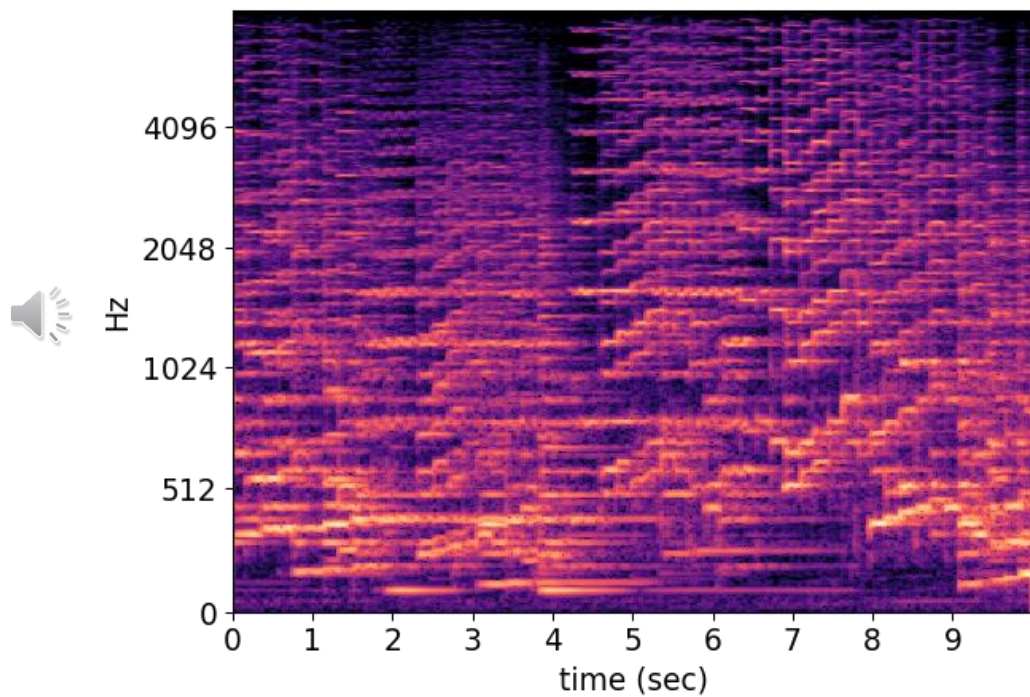
Melodyne Editor



(Source: Celemony)

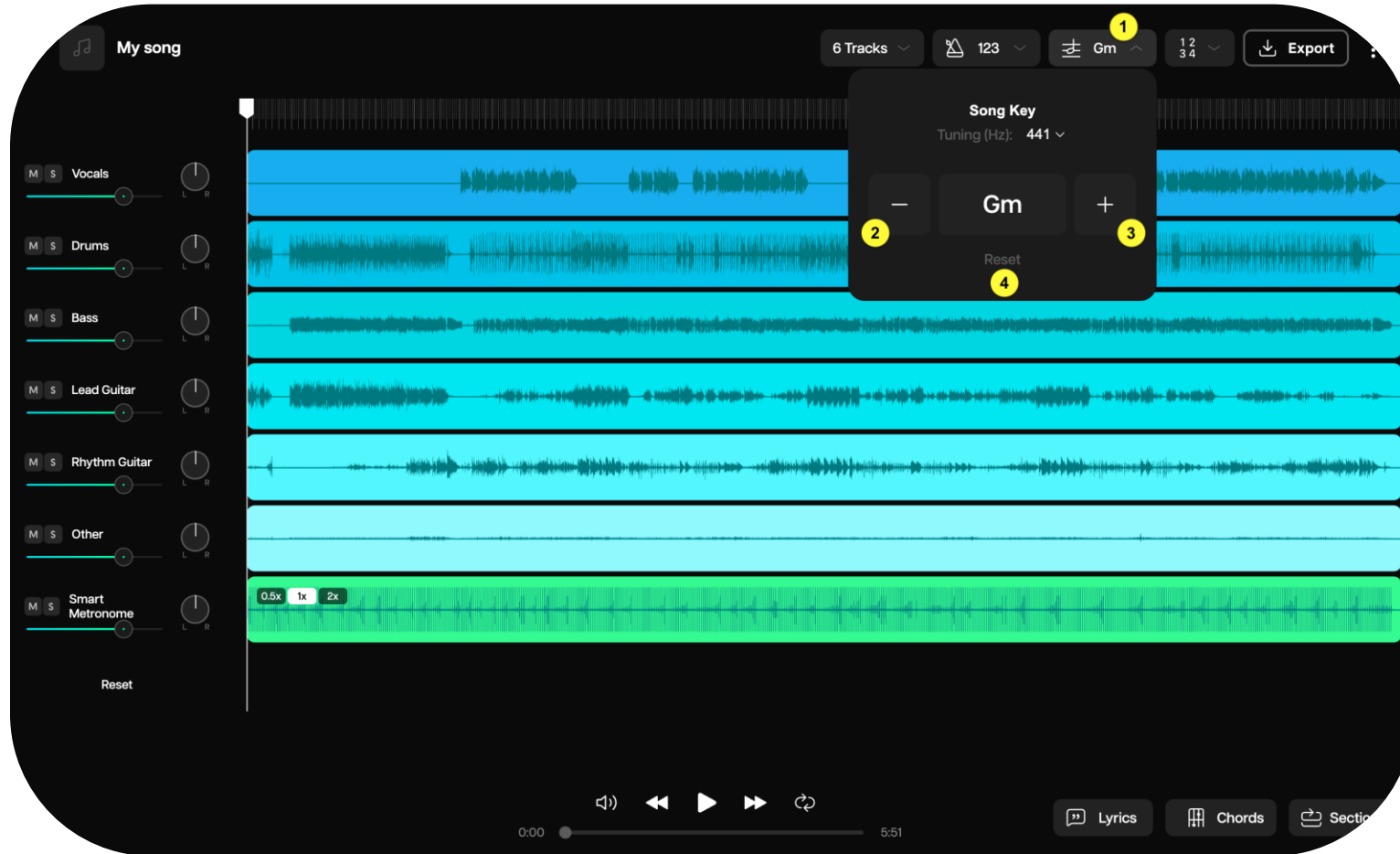
Multitrack Transcription Models

- **MT3** (Gardner et al., 2022)
 - github.com/magenta/mt3



(Source: Gardner et al., 2022)

Key Detection in Moises



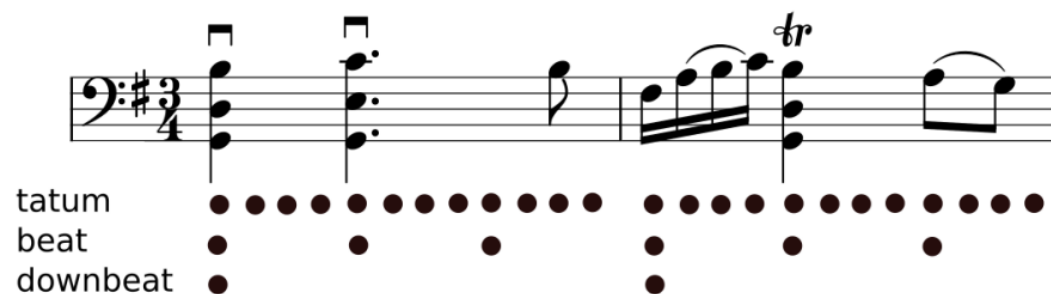
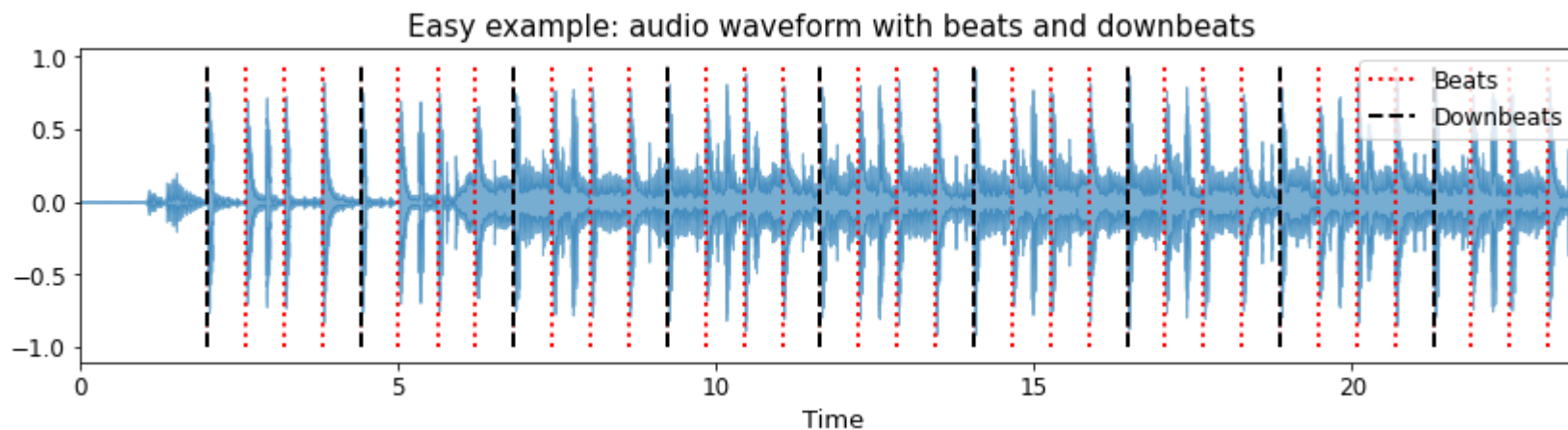
(Source: Moises)

| Chord Detection



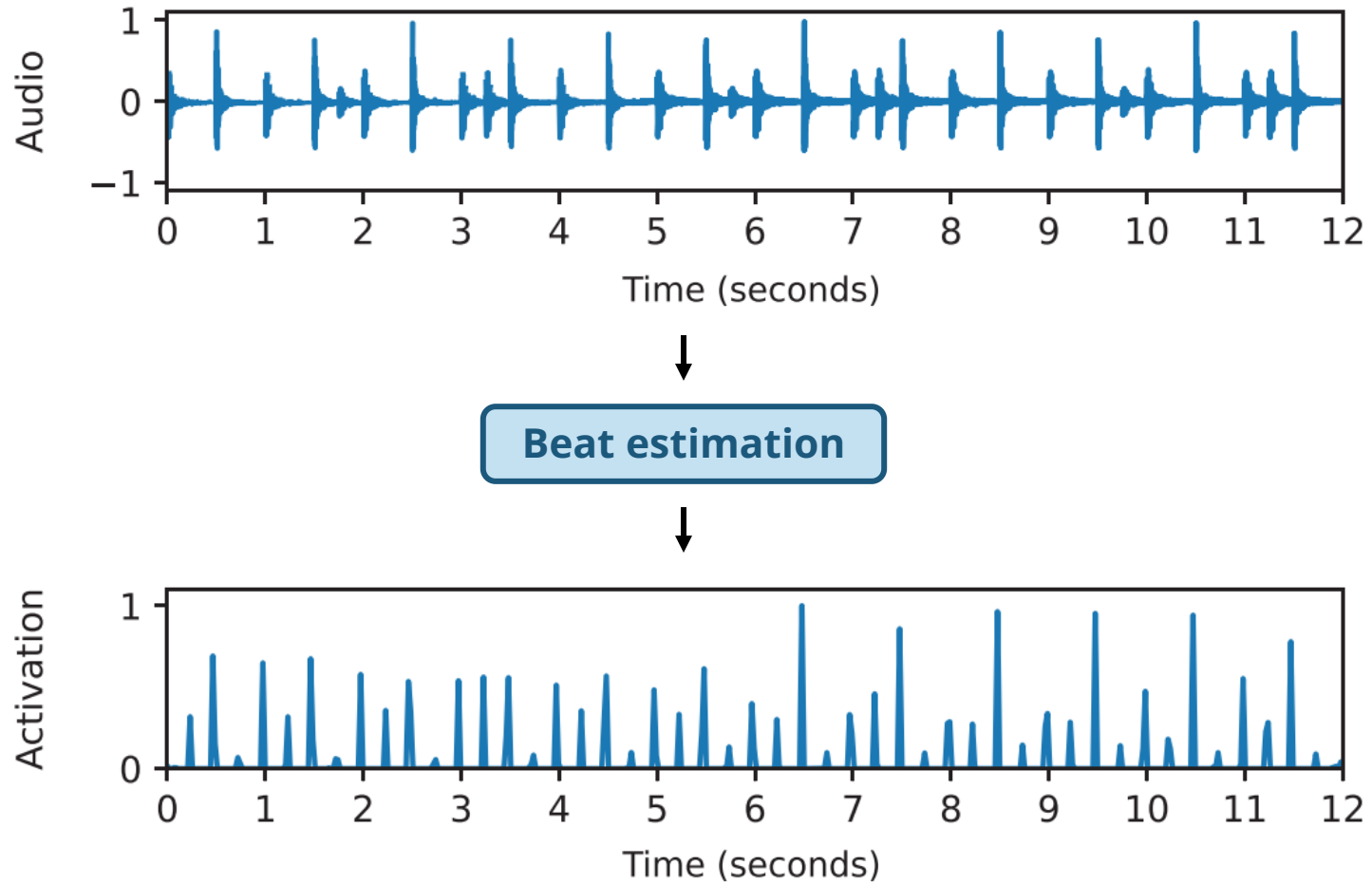
youtube.com/shorts/_N3b_GARMfA

Beat & Downbeat Estimation



(Source: Davies et al., 2021)

Beat & Downbeat Estimation



(Source: Meier et al., 2024)

Structure Analysis

Music segmentation

Allegro.

f marc.

Vivace.

poco ritard.

ff marcato

p poco ritard

A1

A2

B1 B2

C

A3

B3 B4

D

Figure 4.5 following [Müller, FMP, Springer 2015]

(Source: Müller & Zalkow, 2019)

Hierarchical music segmentation

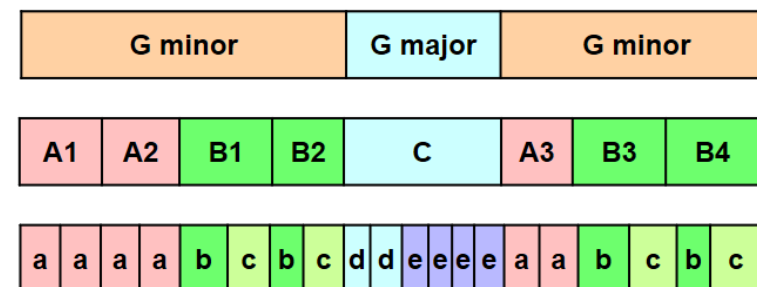
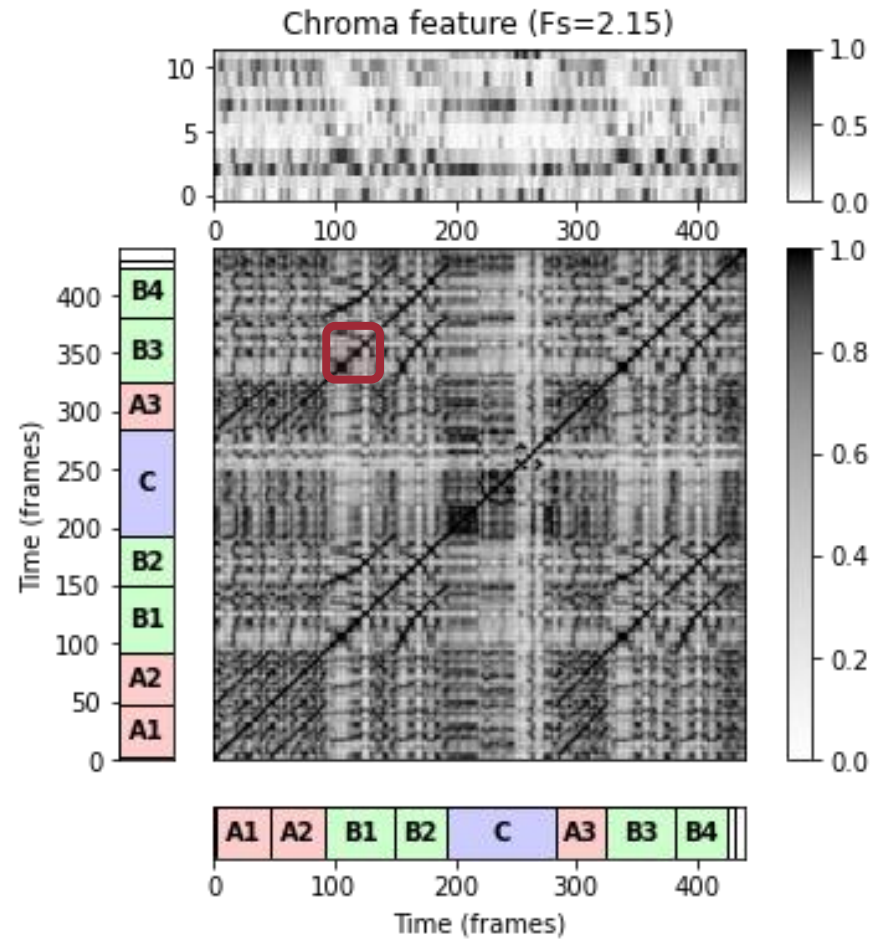


Figure 4.28 from [Müller, FMP, Springer 2015]

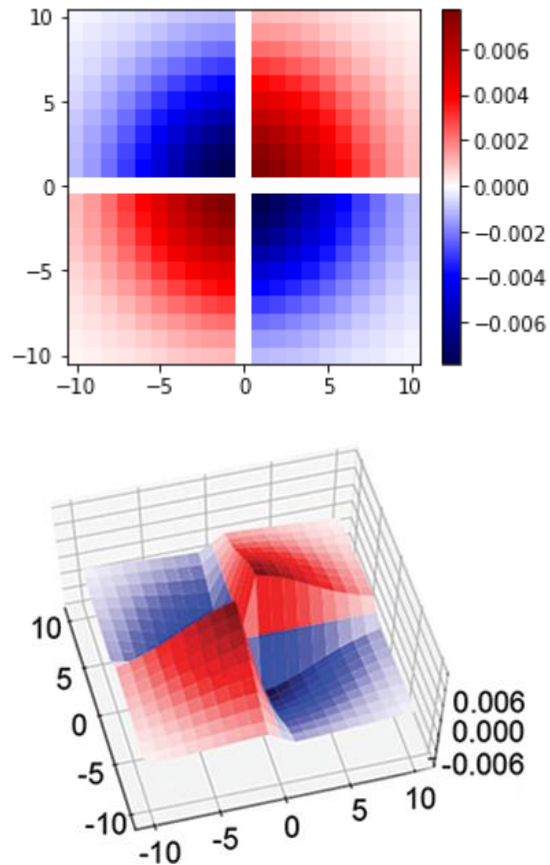
(Source: Müller & Zalkow, 2019)

Self-Similarity Matrices (SSMs)



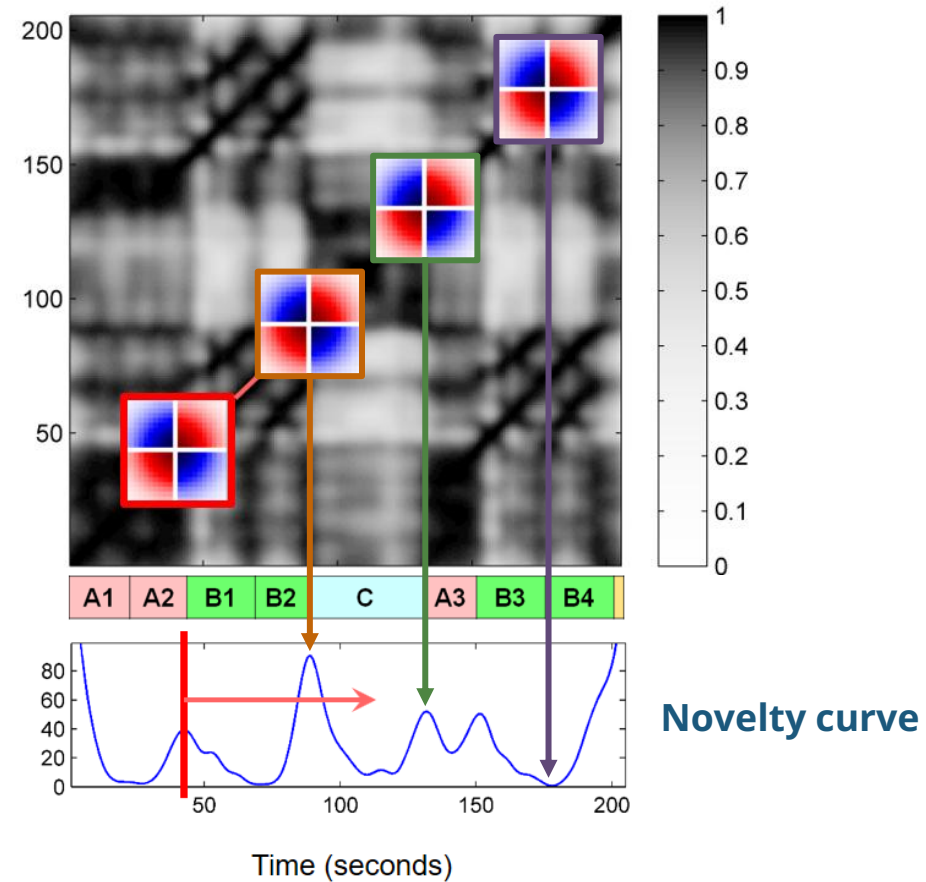
(Source: Müller & Zalkow, 2019)

Self-Similarity Matrices (SSMs)



(Source: Müller & Chiu, 2024)

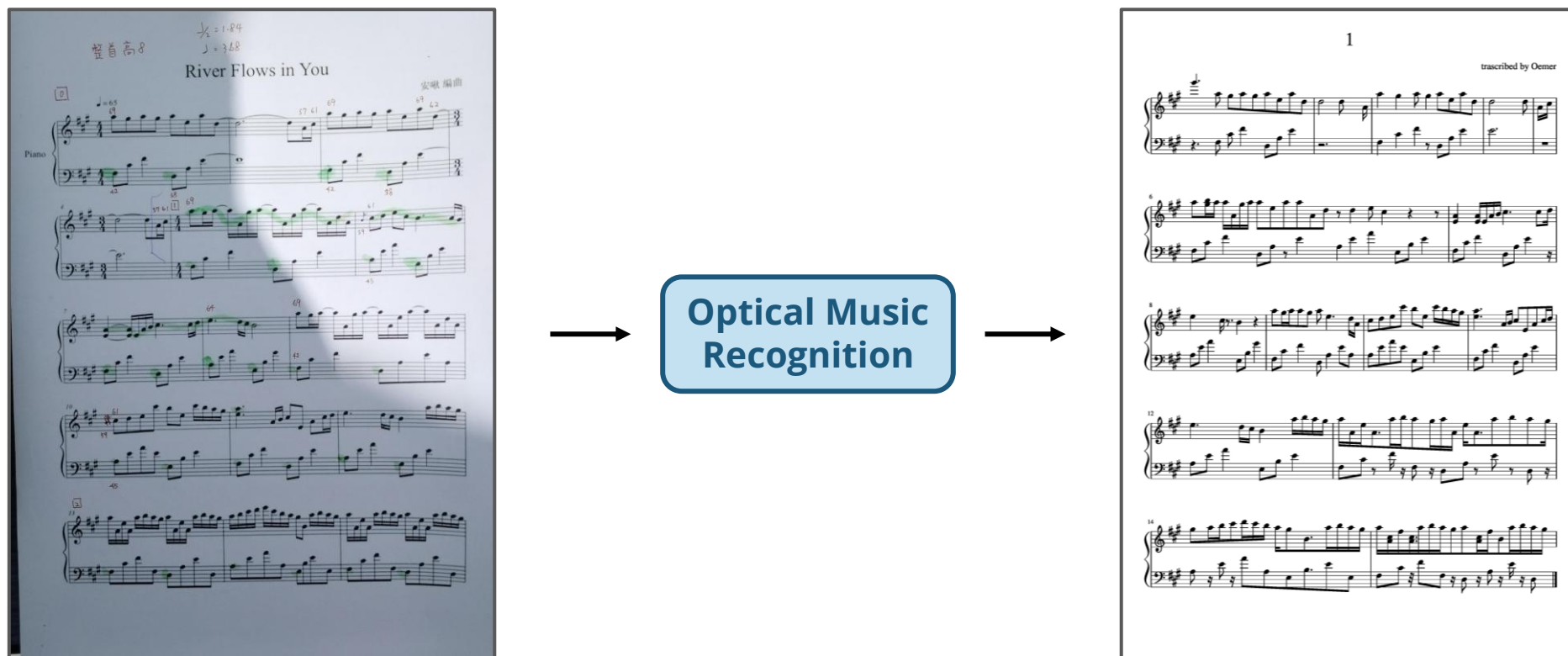
Figure 4.24 from [Müller, FMP, Springer 2015]



(Source: Müller & Zalkow, 2019)

Optical Music Recognition (OMR)

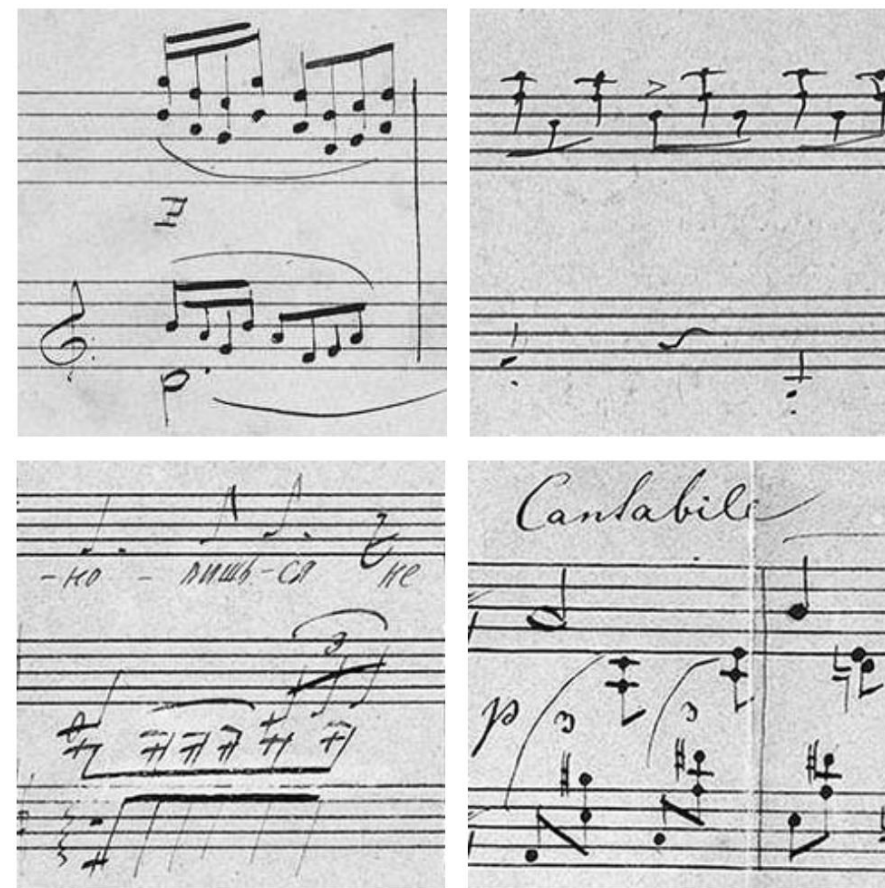
- **Goal:** Convert **scanned sheet music** into **digital musical notation**



Challenges of OMR



(Source: Calvo-Zaragoza et al., 2018)

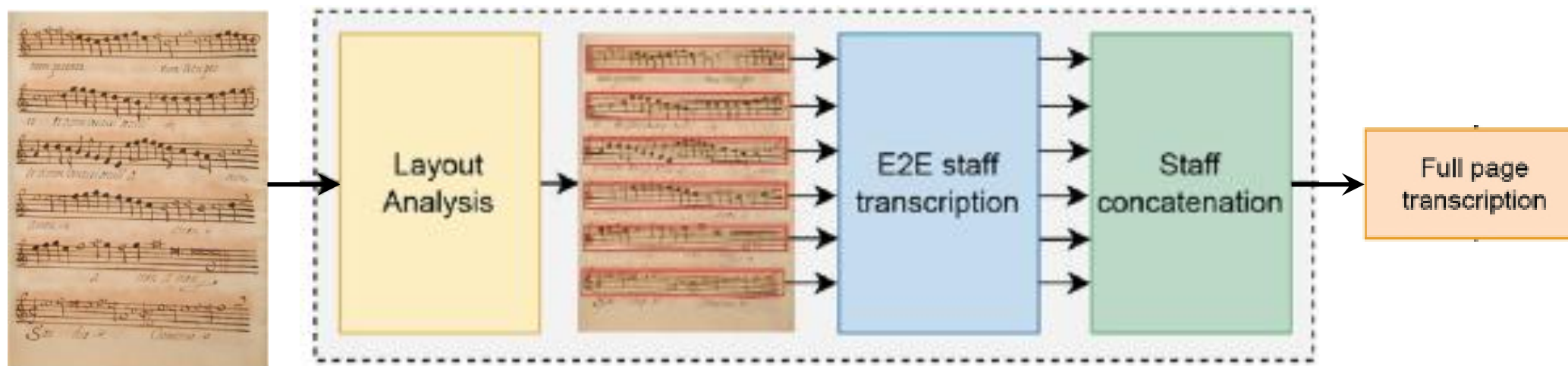


(Source: Novotný & Pokorný, 2015)

Jorge Calvo-Zaragoza, Juan C. Martinez-Sevilla, Carlos Penarrubia, and Antonio Rios-Vila, "Optical Music Recognition: Recent Advances, Current Challenges, and Future Directions," *ICDAR*, 2023.

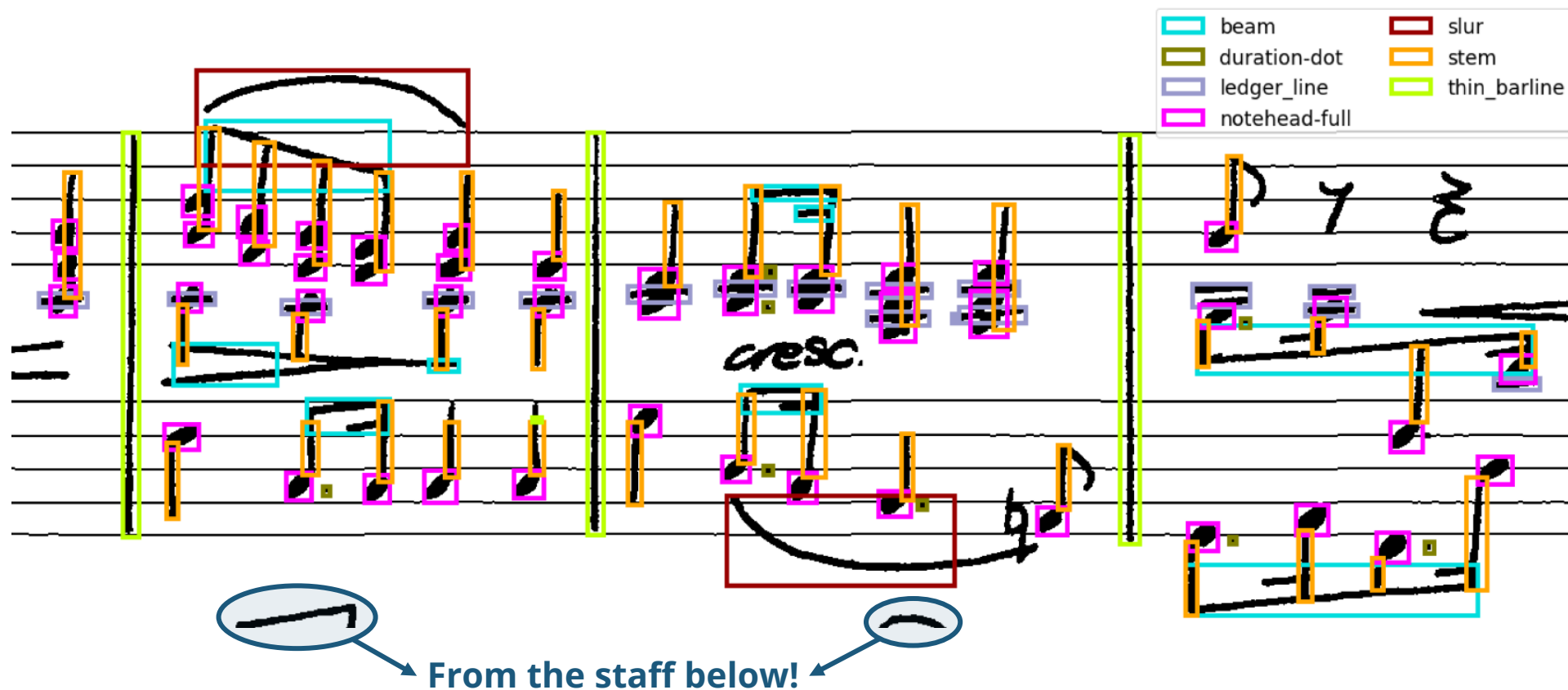
Jiří Novotný and Jaroslav Pokorný, "Introduction to Optical Music Recognition: Overview and Practical Challenges," *DATESO*, 2015.

Common Pipeline of OMR Systems



(Source: Calvo-Zaragoza et al., 2018)

Musical Object Recognition



(Source: Pacha et al., 2018)

Next Lecture

Music Classification

