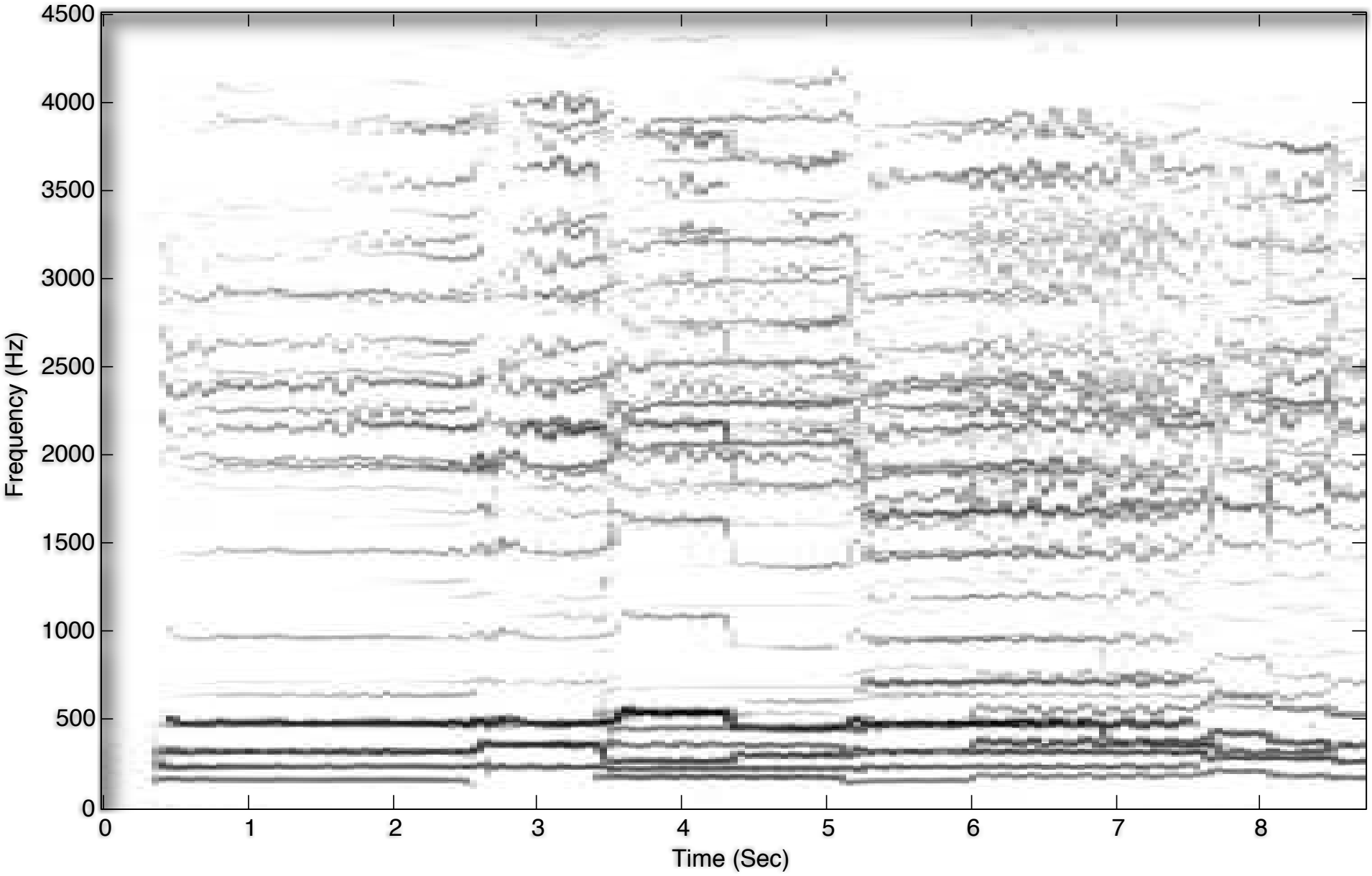
Fundamental Frequency Estimation in Music Tutorial

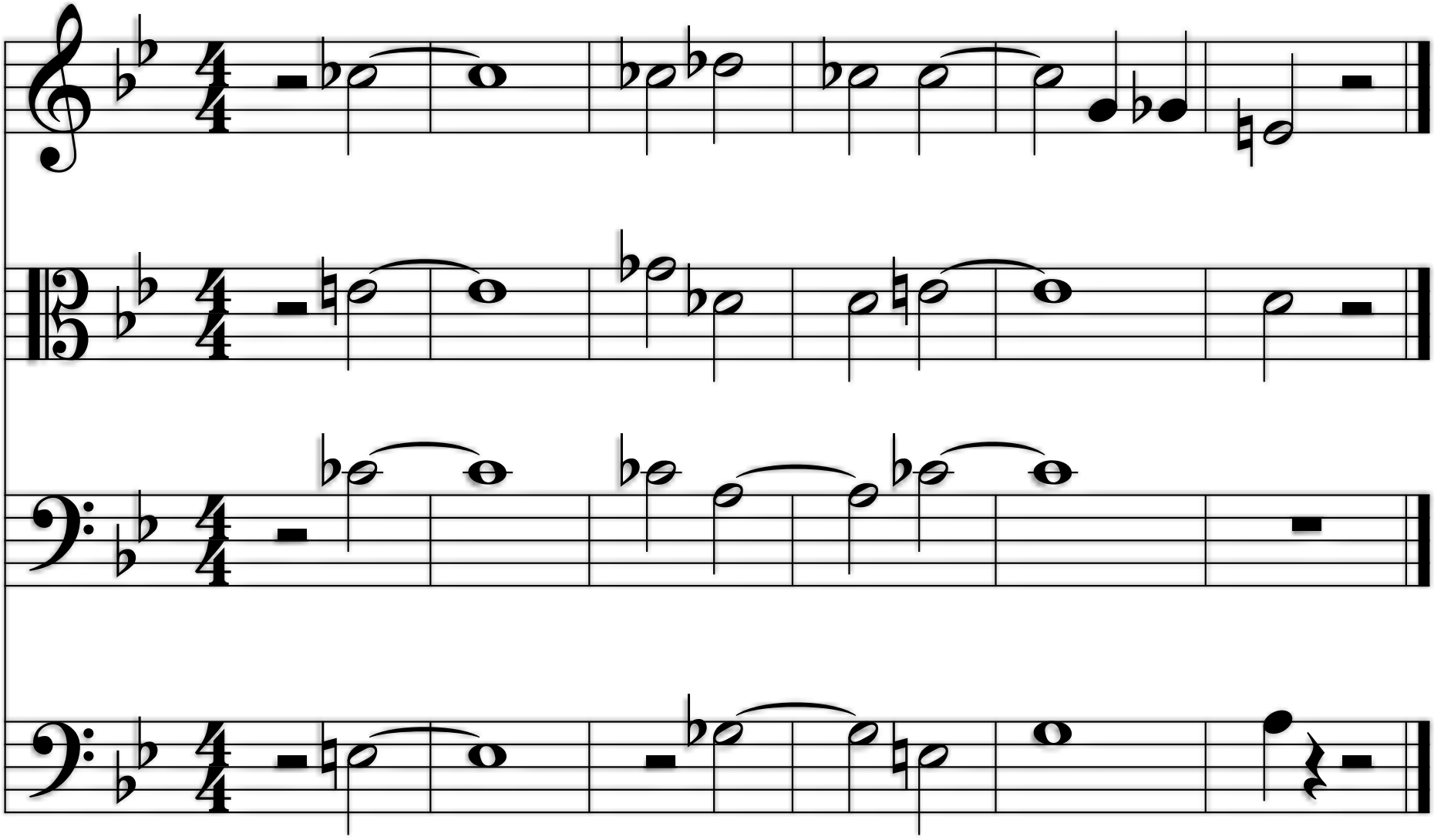
Part 3: Applications

Johanna Devaney
Brooklyn College, CUNY

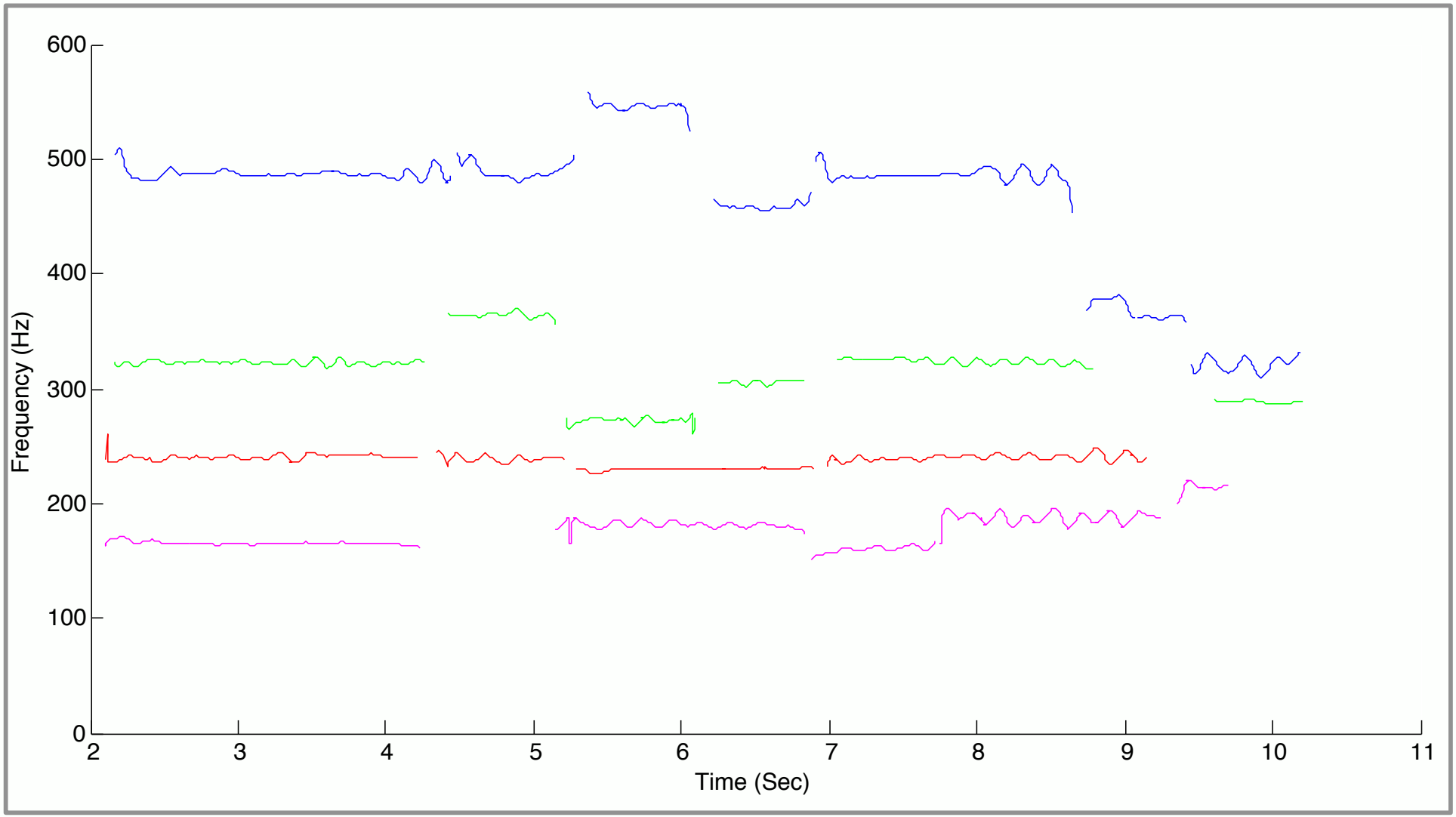
johanna.devaney@brooklyn.cuny.edu


4500
4000
3500
3000
2500
2000
1500
1000
500
0
Frequency (Hz)
0
1
2
3
4
5
6
7
8
Time (Sec)

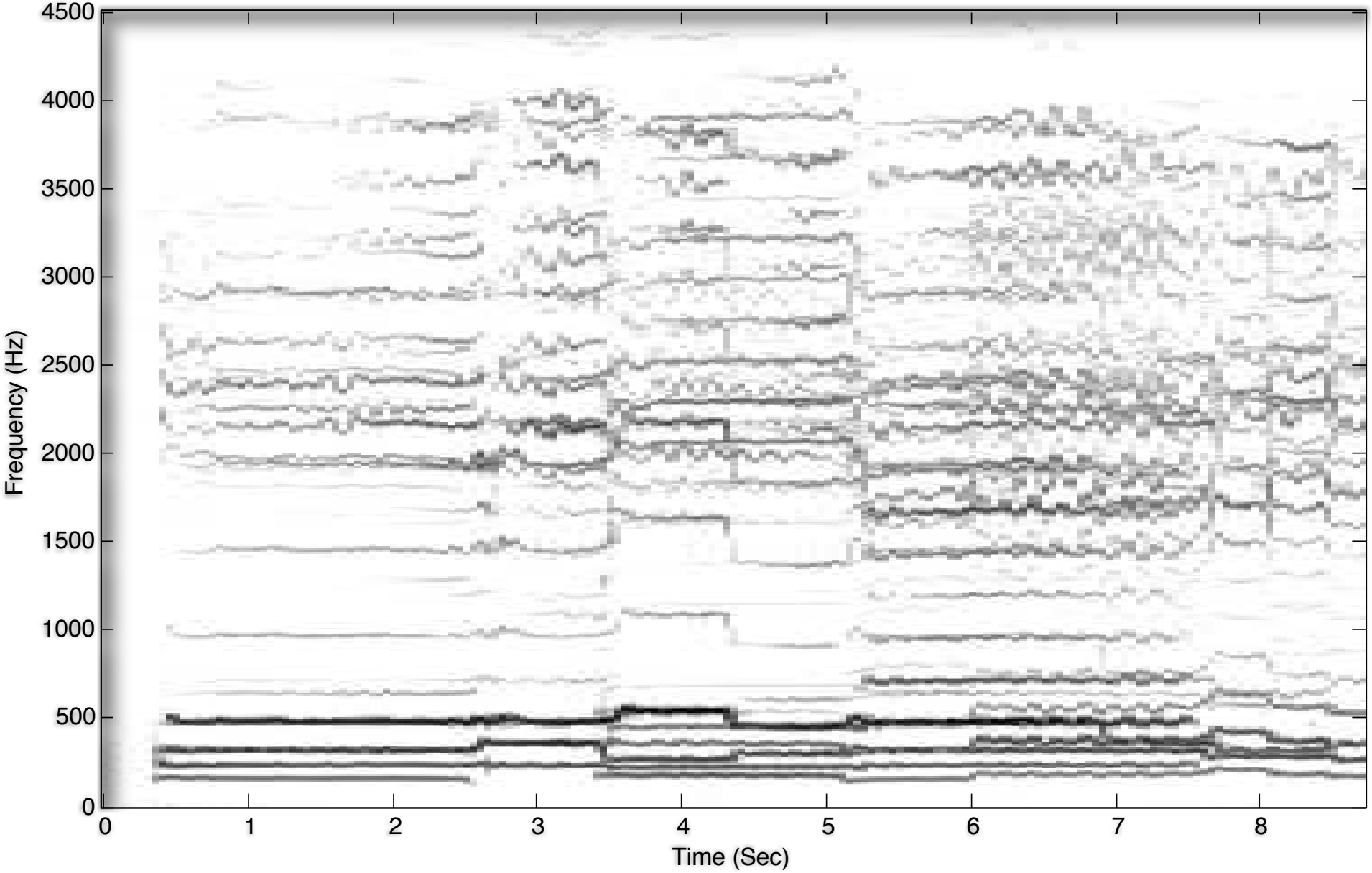
Music Transcription



F_0 Trajectories



Source Separation



Transcription

Automatic Music Transcription

- **Automatic music transcription aims to mimic musicians' ability to transcribe audio into a notated score**
- **multiple f_0 detection is only one subtask of automatic music transcription**
  - Other tasks include (Benetos et al. 2013):
    - note onset/offset detection
    - extraction of rhythmic information and time quantization
    - loudness estimation and quantization
    - instrument recognition

Transcription

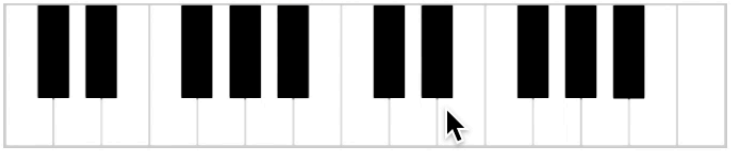
Searching Musical Data

- **Full transcription would allow for audio files to be indexed and searched in a manner similar to musical scores**
- **e.g., an audio version of PeachNote's N-Gram viewer (Viro 2011), which allows users to search for melodies or chords in a number of music libraries**
  - including Petrucci (which has over nearly 450,000 musical scores
  - compare this to Spotify's ~30 million audio tracks

Music Ngram Viewer

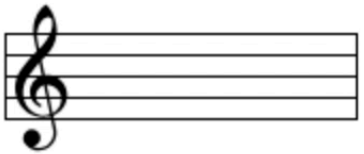
Tweet

Please enter a melody or a sequence of chords (advanced use)



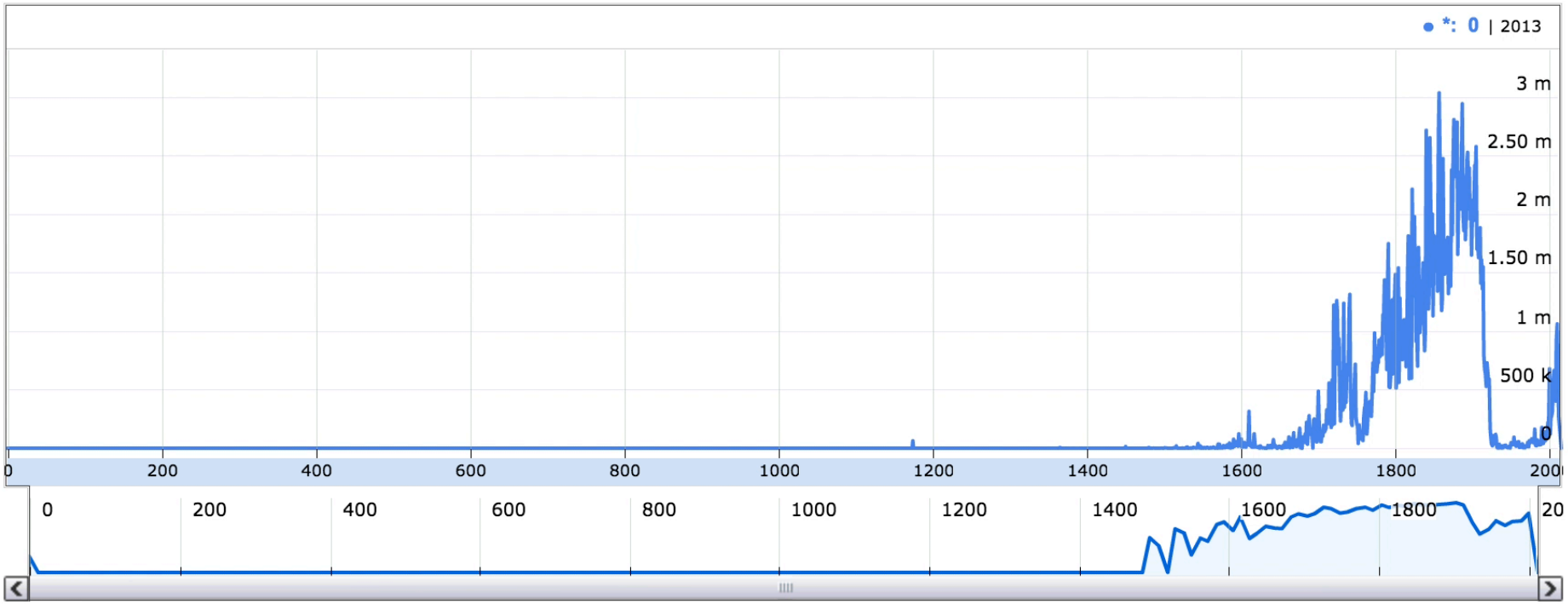
melody Search

Petrucci Music Library Smoothing: 0 Normalized



Keyboard shortcuts

Showing the number of notes indexed for each year.



Run your own experiment! Raw data is available for download here.

Feedback

Transcription

Similarity

- **Score-based similarity approaches are largely focused on Western music**
- **Audio-based methods have traditionally been based on timbral features rather than melody and harmony**
- **Audio-based data would allow for a wider range of musical languages to be modeled**

Source Separation

Interactive Music Systems

- **Improved score-following**
  - Single-instrument score-following is largely a solved problem which has produced usable products (e.g., Music Plus One and Metronaut by Antescofo)
  - Robust multiple f0 estimation could improve single-instrument following in multiple-instrument audio and multi-instrument following

Metronaut by Antescofo

Transcription

Musicology

- **Transcription facilitates the analysis of un-notated music**
  - Particularly music outside of the Western art music tradition
- **Long history of musicological score-based corpus studies that goes back to 1960s**
  - Current work is typically done in Humdrum or music21
  - Translating audio into score notation would allow musicologists to use existing tools for analysis

Introduction 1
Overview

Transcription-Related 2
Music Search, Music Similarity, Interactive Music Systems, Musicology

Performance-Related 3
Note-level Descriptors, Music Generation, Style Transfer

Source Separation-Related 4
Melody Estimation, Music Retrieval, Music Production

Conclusions 5
Summary

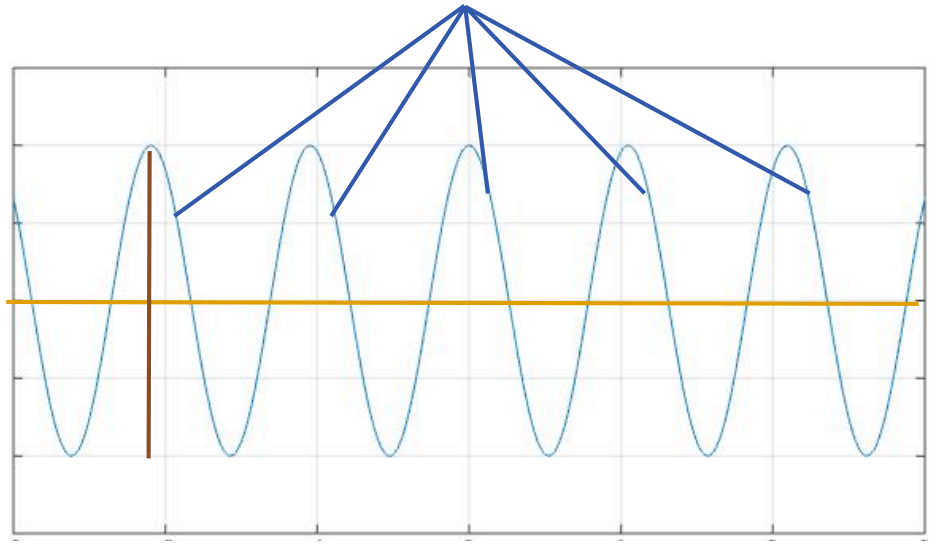
Performance

Note-level descriptors

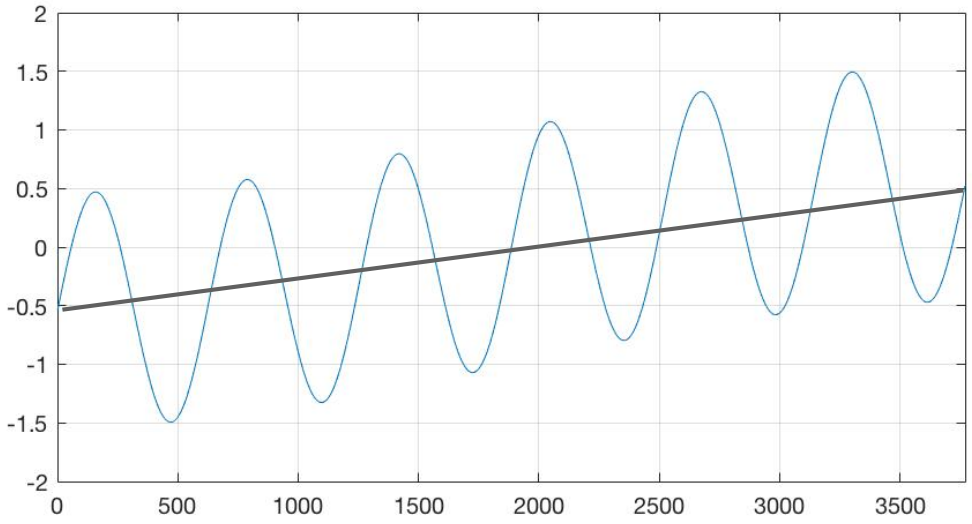
Perceived pitch

Vibrato Depth

Vibrato Rate



Slope



Performance

Note-level descriptors

- **Perceived Pitch**
  - Based on psychophysical testing
  - Simple mean, e.g., arithmetic or geometric (Shonle and Horan 1980)
  - Weighted mean, e.g., based on f_0 rate of change (Gockel et al. 2001)

Performance

Note-level descriptors

- **Vibrato**
  - Fitting a sinusoid to a normalized version of the f_0 trace (Panteli et al. 2017)
  - Applying an FFT, the position of second peak is the vibrato rate and its value is the depth (Devaney 2015)

Performance

Note-level descriptors

- **Slope (and Curvature)**
  - Fitting a polynomial to f_0 trace (Panteli et al. 2017)
  - Decomposing f_0 trace with the DCT (Devaney et al. 2011)

Performance

Manual and Score-Informed Approaches

- **Much of the work on note-level performance analysis has either been done**
  - Manually using tools like Sonic Visualiser or TONY
  - With score-informed approaches
    - Turkish makam (Şentürk et al. 2014)
    - Jazz solos (Abeßer et al. 2017)
    - Classical string performances (Li et al. 2017)
    - Classical vocal performances (Devaney and Mandel 2017)

Performance

Contours

- **In the absence of full transcriptions, contours can been used for musicological corpus studies**
- **The same techniques used for note-level descriptors can be applied to contours**
- **Shown to be useful for characterizing singing style (Panteli et al. 2017)**

Performance

Music Generation

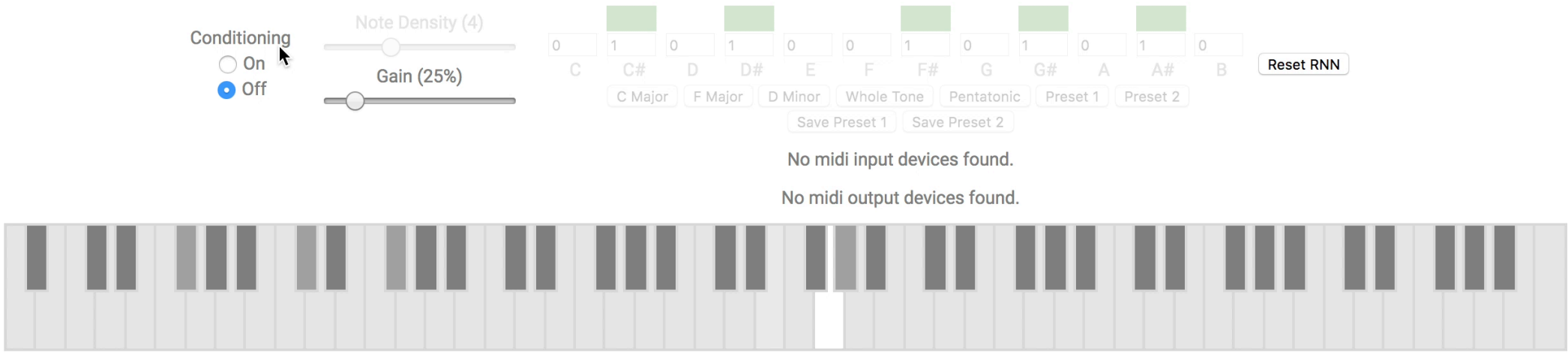
- **Automatic music generation greatly benefits from automatic transcription, but the task also requires performance data in order to produce human-sounding performances**
- **E.g., an audio trained version of Magenta's Performance RNN, which uses MIDI recordings from the Yamaha ePiano Competition**

Quantized

Human performance

Audio examples from magenta.tensorflow.org

Performance RNN



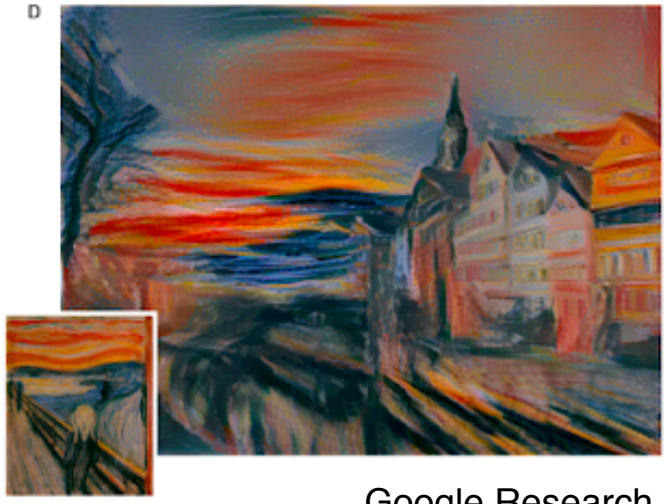
magenta

[Performance RNN](#) was trained in TensorFlow on MIDI from piano performances. It was then ported to run in the browser using only Javascript in the [TensorFlow.js](#) environment. Piano samples are from [Salamander Grand Piano](#).

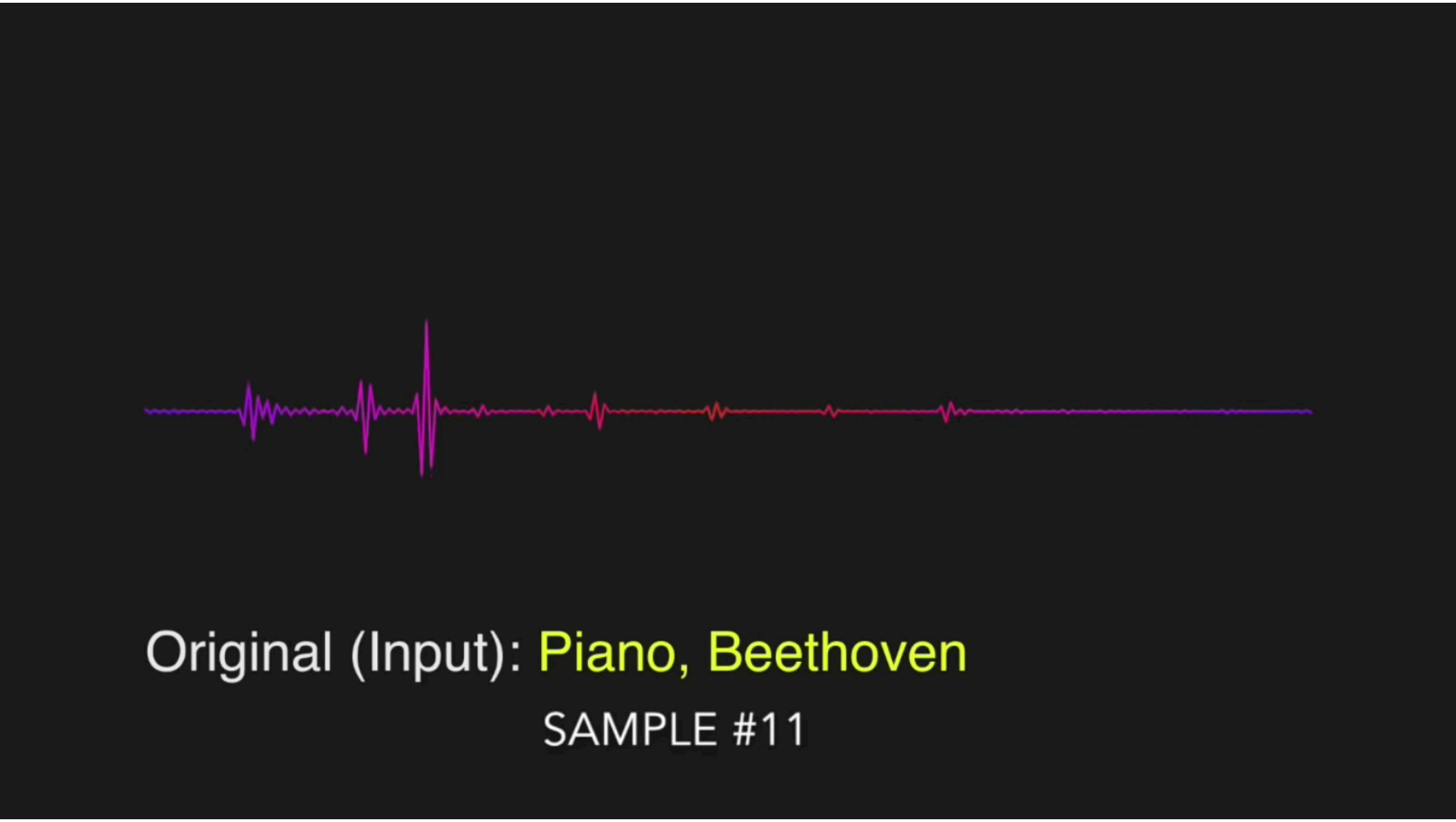
magenta.tensorflow.org

Performance

Style Transfer



Google Research



Facebook Artificial Intelligence Research

Introduction 1
Overview

Transcription-Related 2
Music Search, Music Similarity, Interactive Music Systems, Musicology

Performance-Related 3
Note-level Descriptors, Music Generation, Style Transfer

Source Separation-Related 4
Melody Estimation, Music Retrieval, Music Production

Conclusions 5
Summary

Source Separation

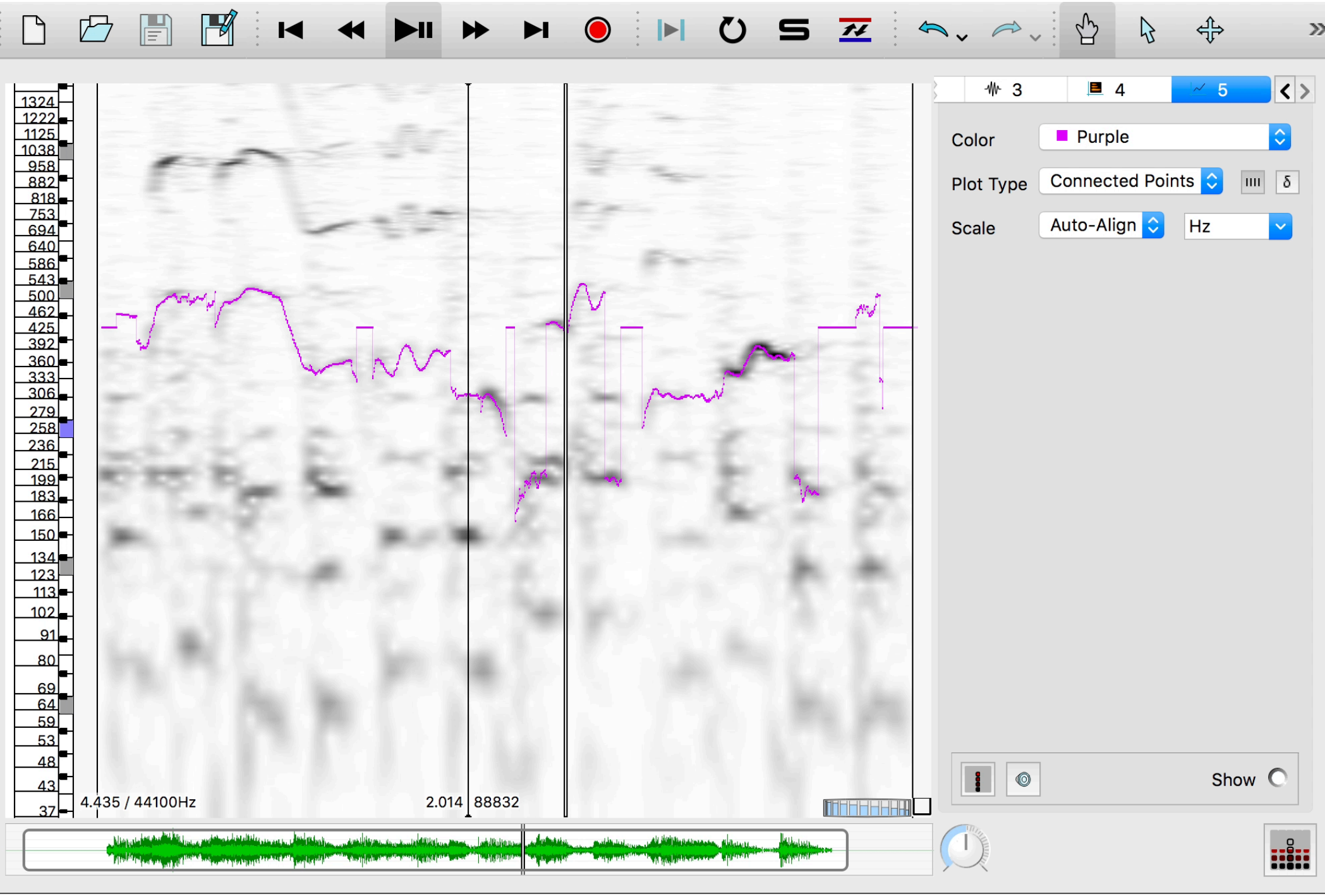
Overview

- **The practical relationship between multiple f_0 estimation and source separation is not entirely clear**
  - Multiple f_0 estimation could inform source separation
  - Source separation could inform multiple f_0 estimation
  - Joint estimation
- **Much work remains to be done in this area and currently many of the solutions are score-informed**

Source Separation

Melody Extraction

- **Melody extraction algorithms estimate f_0 for the dominant melodic line in a polyphonic musical recording**
- **Active area of research in MIR (Salamon et al. 2014)**
- **Related to bass and other single-line transcription tasks**



Melodia VAMP Plugin

Source Separation

Music Retrieval

- **Ability to extract the melody facilitates a number of retrieval tasks**
  - Query by humming - where the user sings or hums a melody and the system finds matching examples of recorded audio
  - Query by example - where the user provides an audio recording with a melody to be matched the system finds matching examples of recorded audio
    - A specific case of query-by-example is cover song identification
      - Particularly when the melody is preserved in the cover but other musical aspects are varied

Source Separation

Music Retrieval

- **Extraction of individual musical lines (melody or otherwise) can facilitate classification tasks**
  - Performer classification - *as with the work discussed earlier using contours for characterizing singing*
  - Genre classification - *certain performance styles are characteristic of different genre styles*
    - as is the balance between the vocals and instruments
    - this can provide valuable data in addition to analyses of the full audio signal

Source Separation

Music Production

- **Full separation of the music sources in a signal can facilitate a number of music production tasks**
  - Generating new mixes - *where the relative levels of the sources in a pre-mixed audio source can be adjusted*
  - De-soloing - *where the lead melodic line can be removed*
    - Useful for
      - resetting the lead melodic line against other musical material
      - karaoke applications using only the accompaniment

Source Separation

Music Production

- **Multiple f_0 detection is also useful for guiding audio processing algorithms**
  - e.g., retuning vocals

Melodyne

Autotone

Images from soundonsound.com

Introduction 1
Overview

Transcription-Related 2
Music Search, Music Similarity, Interactive Music Systems, Musicology

Performance-Related 3
Note-level Descriptors, Music Generation, Style Transfer

Source Separation-Related 4
Melody Estimation, Music Retrieval, Music Production

Conclusions 5
Summary

Conclusions

Summary

- **Transcription-Related**
  - Music Search, Music Similarity, Interactive Music Systems, Musicology
- **Performance-Related**
  - Note-level Descriptors, Music Generation, Style Transfer
- **Source Separation-Related**
  - Melody Estimation, Music Retrieval, Music Production

Thank you!

Questions?

References

Abeßer, Jakob, Klaus Frieler, Estefanía Cano, Martin Pfleiderer, Wolf-Georg Zaddach. "Score-informed analysis of tuning, intonation, pitch modulation, and dynamics in jazz solos." *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)* 25, no. 1 (2017): 168-177.

Benetos, Emmanouil, Simon Dixon, Dimitrios Giannoulis, Holger Kirchhoff, and Anssi Klapuri. "Automatic music transcription: challenges and future directions." *Journal of Intelligent Information Systems* 41, no. 3 (2013): 407-434.

Devaney, Johanna C., Michael I. Mandel, and Ichiro Fujinaga. "Characterizing singing voice fundamental frequency trajectories." In *Proceedings of Workshop on Applications of Signal Processing to Audio and Acoustics* (WASPAA), 2011 IEEE Workshop on, pp. 73-76. IEEE, 2011.

Devaney, Johanna. "Recapturing the musical performance data in Seashore's published performance scores." *Musicae Scientiae* 19, no. 2 (2015): 214-222.

Devaney, Johanna, and Michael Mandel. "An evaluation of score-informed methods for estimating fundamental frequency and power from polyphonic audio." In *Proceedings of International Conference on Acoustics, Speech and Signal Processing* (ICASSP), 2017, pp. 181-185. IEEE, 2017.

Gockel, Hedwig, Brian CJ Moore, and Robert P. Carlyon. "Influence of rate of change of frequency on the overall pitch of frequency-modulated tones." *Journal of the Acoustical Society of America* 109, no. 2 (2001): 701-712.

References

Li, Bochen, Karthik Dinesh, Zhiyao Duan, and Gaurav Sharma. "See and listen: Score-informed association of sound tracks to players in chamber music performance videos." In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (ICASSP), pp. 2906-2910. IEEE, 2017.

Panteli, Maria, Rachel Bittner, Juan Pablo Bello, and Simon Dixon. "Towards the characterization of singing styles in world music." In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 636-640. IEEE, 2017.

Salamon, Justin, Emilia Gómez, Daniel PW Ellis, and Gaël Richard. "Melody extraction from polyphonic music signals: Approaches, applications, and challenges." *IEEE Signal Processing Magazine* 31, no. 2 (2014): 118-134.

Şentürk, Sertan, Andre Holzapfel, and Xavier Serra. "Linking scores and audio recordings in makam music of Turkey." *Journal of New Music Research* 43, no. 1 (2014): 34-52.

Shonle, John I., and Kathryn E. Horan. "The pitch of vibrato tones." *Journal of the Acoustical Society of America* 67, no. 1 (1980): 246-252.

Viro, Vladimir. "Peachnote: Music Score Search and Analysis Platform." In *Proceedings of International Society for Music Information Retrieval* conference, pp. 359-362. ISMIR, 2011.