Introductions to Music Information Retrieval

ECE 272/472 Audio Signal Processing

Bochen Li

University of Rochester

Wish List

For music learners/performers

- While I play the piano, turn the page for me
- Tell me if I play wrong notes
- While I sing a song, automatic play accompaniment for me

Wish List

For concert audiences

- Tell me what is the instrument being played from orchestra
- Tell me what is the pitch/chord/key/tempo of the music being played
- Display the lyrics for a choir performance
- Just record/play the solo part, and mute the others

Wish List

For musicologists

- Transcribe an improvised piano jazz performance into sheet music
- Numerically compare different artists' performances expressiveness and personal styles
- Scan a sheet music into computer readable format, e.g., MIDI, XML, Lilypond, etc.
- Generate music following a composer' style, e.g., bring *Chopin* back to life

Wish List

For music listeners

- Always play my favorite song in radio stream
- Sing a song fragment and find out the name
- Automatically play the song for my mood

Introduction

What is Music Information Retrieval (MIR)?

- Audio Signal Processing
- Machine Learning
- Musicology
- Psychoacoustics
- Computer Vision

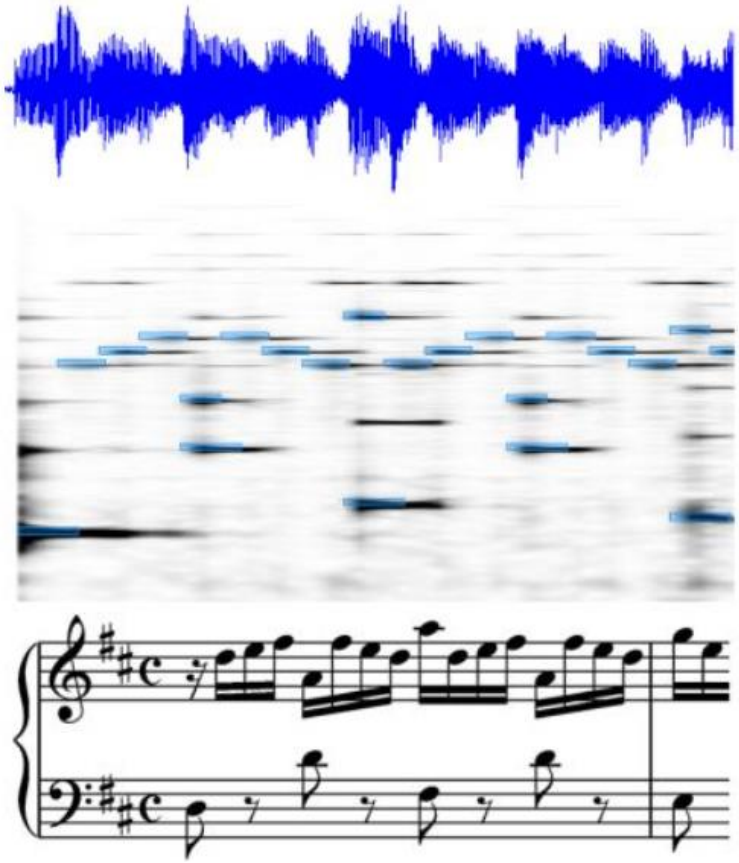
Tasks

- **Automatic Music Transcription**
- Music Synchronization
- Source Separation
- Performance Expressiveness Analysis

Tasks

Automatic Music Transcription

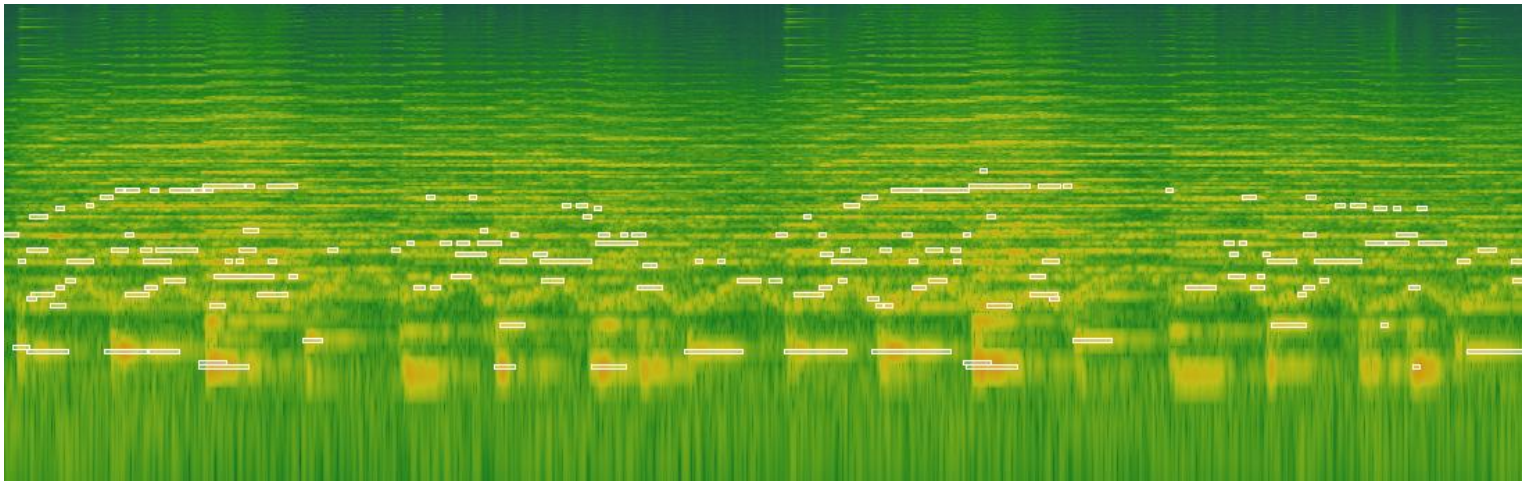
the process of converting an acoustic musical signal into some form of music notation (e.g. staff notation, MIDI file, piano-roll,...)



Tasks

Automatic Music Transcription

the process of converting an acoustic musical signal into some form of music notation (e.g. staff notation, MIDI file, piano-roll,...)







Tasks

Automatic Music Transcription

Subtasks:

- Pitch detection
- Onset/offset detection
- Instrument identification
- Rhythm parsing
- Identification of dynamics/expression

Tasks

Automatic Music Transcription

State-of-the-art Outline:

1. Multi-pitch Analysis
    - Frame-level
    - Note-level
    - Stream-level
2. Towards a Complete Music Notation

Tasks

Automatic Music Transcription

State-of-the-art Outline:

1. Multi-pitch Analysis
   - Frame-level (multi-pitch estimation)
     - Estimate pitches and polyphony in each frame
   - Note-level
     - Estimate pitch, onset, offset of notes
   - Stream-level
     - Stream pitches by sources

Tasks

Automatic Music Transcription

- Frame-level (multi-pitch estimation)

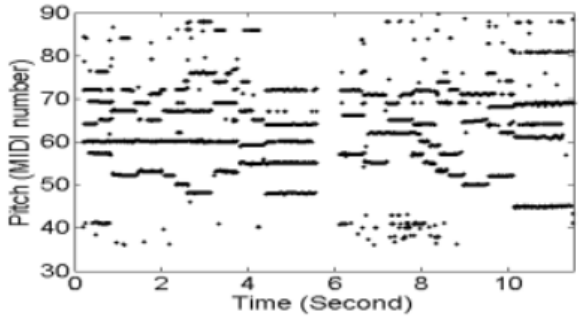
  - Estimate pitches and polyphony in each frame

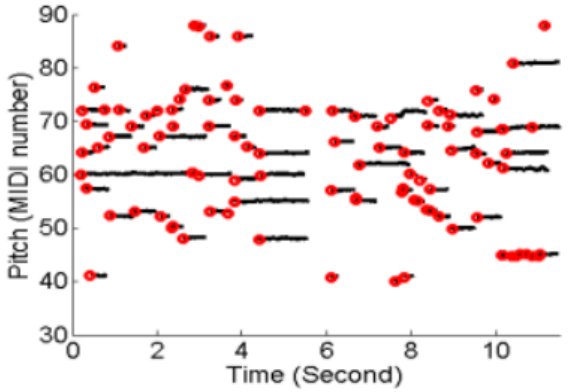
- Note-level

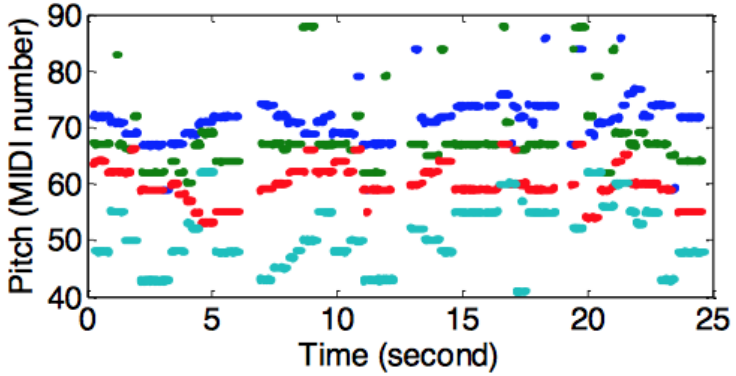
  - Estimate pitch, onset, offset of notes

- Stream-level

  - Stream pitches by sources







Tasks

Automatic Music Transcription

Frame-level (multi-pitch estimation)

Categorization of methods:

- Domain of Operation:
  - Time
  - Frequency
  - Hybrid
- Core Algorithms:
  - Signal processing approaches
  - Maximum likelihood estimation
  - Bayesian
  - Spectrogram decomposition
  - Sparse coding

Tasks

Automatic Music Transcription

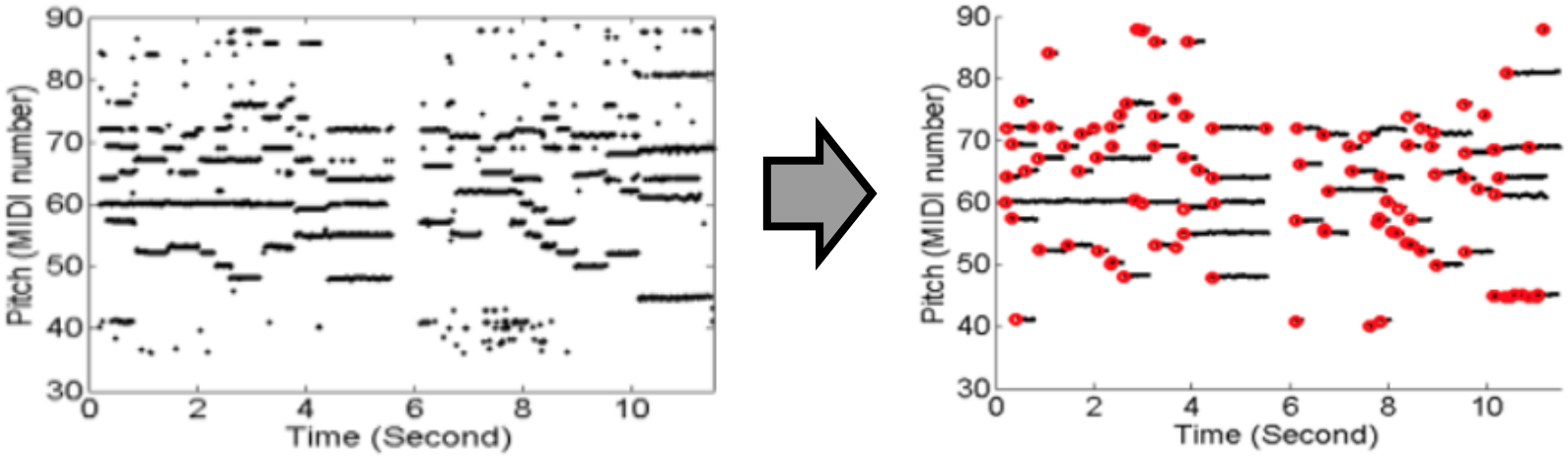
Note-level (Note Tracking)

- Onset Detection

  - Can be sensitive to onset detection accuracy

- Post-processing of frame-level results

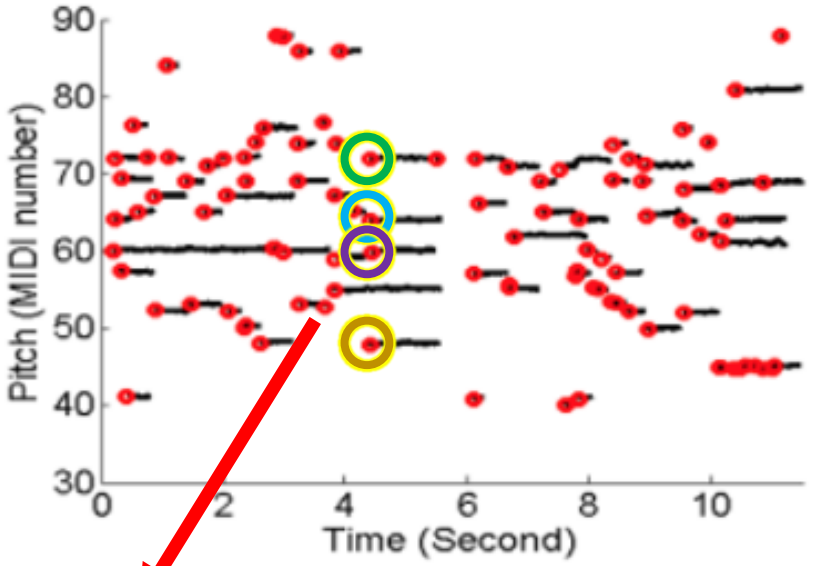
  - Form notes independently by connecting nearby pitches
  - Consider interactions between simultaneous pitches

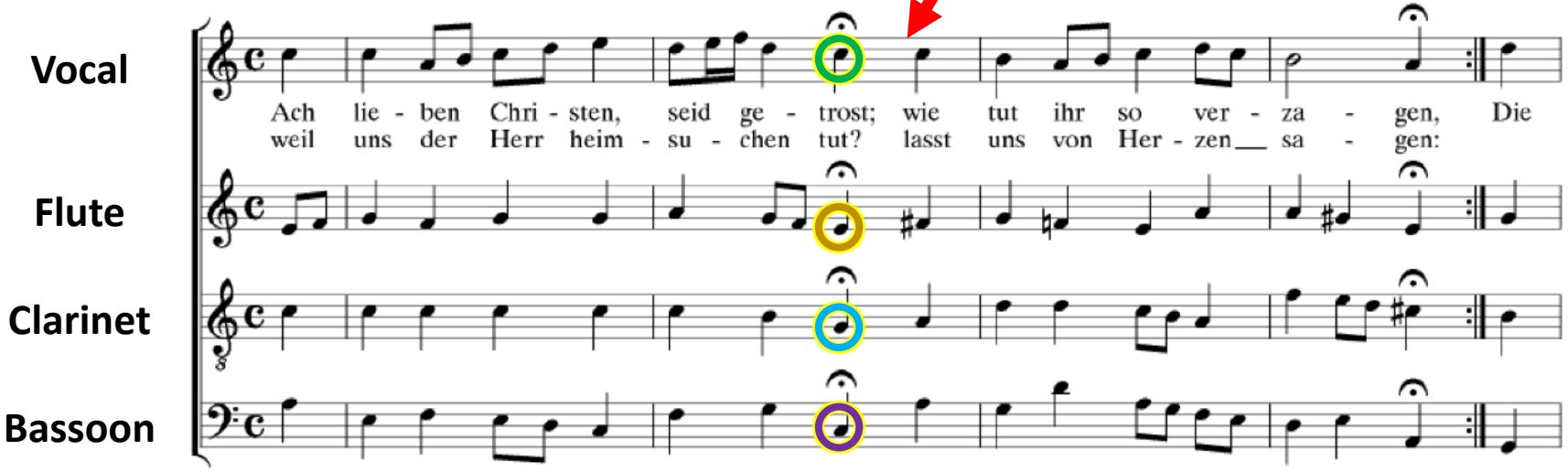


Tasks

Automatic Music Transcription

Stream-level (Timbre-tracking)





Tasks

Automatic Music Transcription

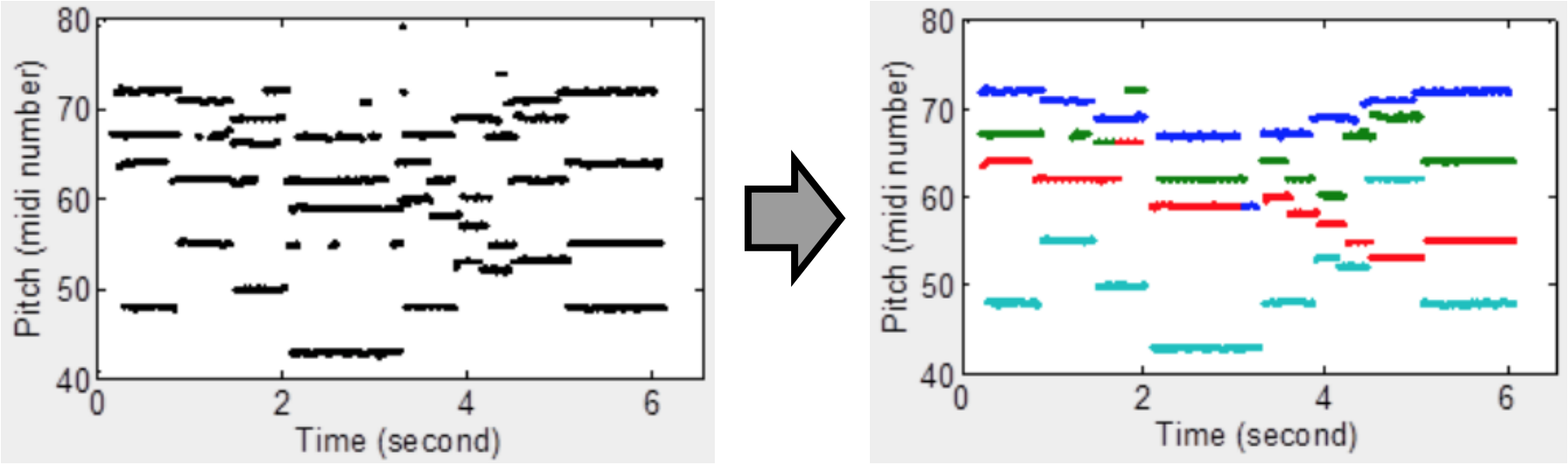
Stream-level (Timbre-tracking)

- Supervised

  - Train timbre models of sound sources

- Unsupervised

  - Cluster pitch estimates according to timbre



Tasks

Automatic Music Transcription

State-of-the-art Results

- Frame-level (Multi-pitch Estimation)

Teams	2009	2010	2011	2012	2013	2014
Yeh and Roebel	0.69	0.69	0.68	-	-	-
Dressler	-	-	0.63	0.64	-	0.68
Canadas-Quesada et al.	-	0.49	-	-	-	-
Benetos and Dixon/Weyde	-	0.47	0.57	0.58	0.66	0.66
Duan et al.	0.57	0.55	-	-	-	-
Fuentes et al.	-	-	-	0.56	-	-
Elowsson and Friberg	-	-	-	-	-	0.72
Cheng et al.	-	-	-	-	0.62	-
Su and Yang	-	-	-	-	-	0.64

Tasks

Automatic Music Transcription

State-of-the-art Results

- Note-level (Note Tracking)

Teams	2009	2010	2011	2012	2013	2014
Yeh and Roebel	0.50	0.53	0.56	-	-	-
Dressler	-	-	-	0.65	-	0.66
Benetos and Dixon/Weyde	-	-	0.45	0.43	0.55	0.58
Duan, Han and Pardo	0.43	0.41	-	-	-	-
Fuentes et al.	-	-	-	0.61	-	-
Elowsson and Friberg	-	-	-	-	-	0.82
Cheng et al.	-	-	-	-	0.50	-
Su and Yang	-	-	-	-	-	0.46
Böck	-	-	-	0.50	-	0.54
Dessein et al.	-	0.40	-	-	-	-
Duan and Temperley	-	-	-	-	-	0.45

Tasks

Automatic Music Transcription

State-of-the-art Results

	Original Audio	Transcription results (played as MIDI synthesis)
Bach's Minuet in G		
Chopin's Etude Op. 10 No. 1		

Tasks

Automatic Music Transcription

State-of-the-art Outline:

1. Multi-pitch Analysis
    - Frame-level
    - Note-level
    - Stream-level
2. Towards a Complete Music Notation

Tasks

Automatic Music Transcription

State-of-the-art Outline:

2. Towards a Complete Music Notation

Current AMT systems can:

- Detect (multiple) pitches, onsets, offsets
- Identify instruments in polyphonic music
- Assign detected notes to a specific instrument

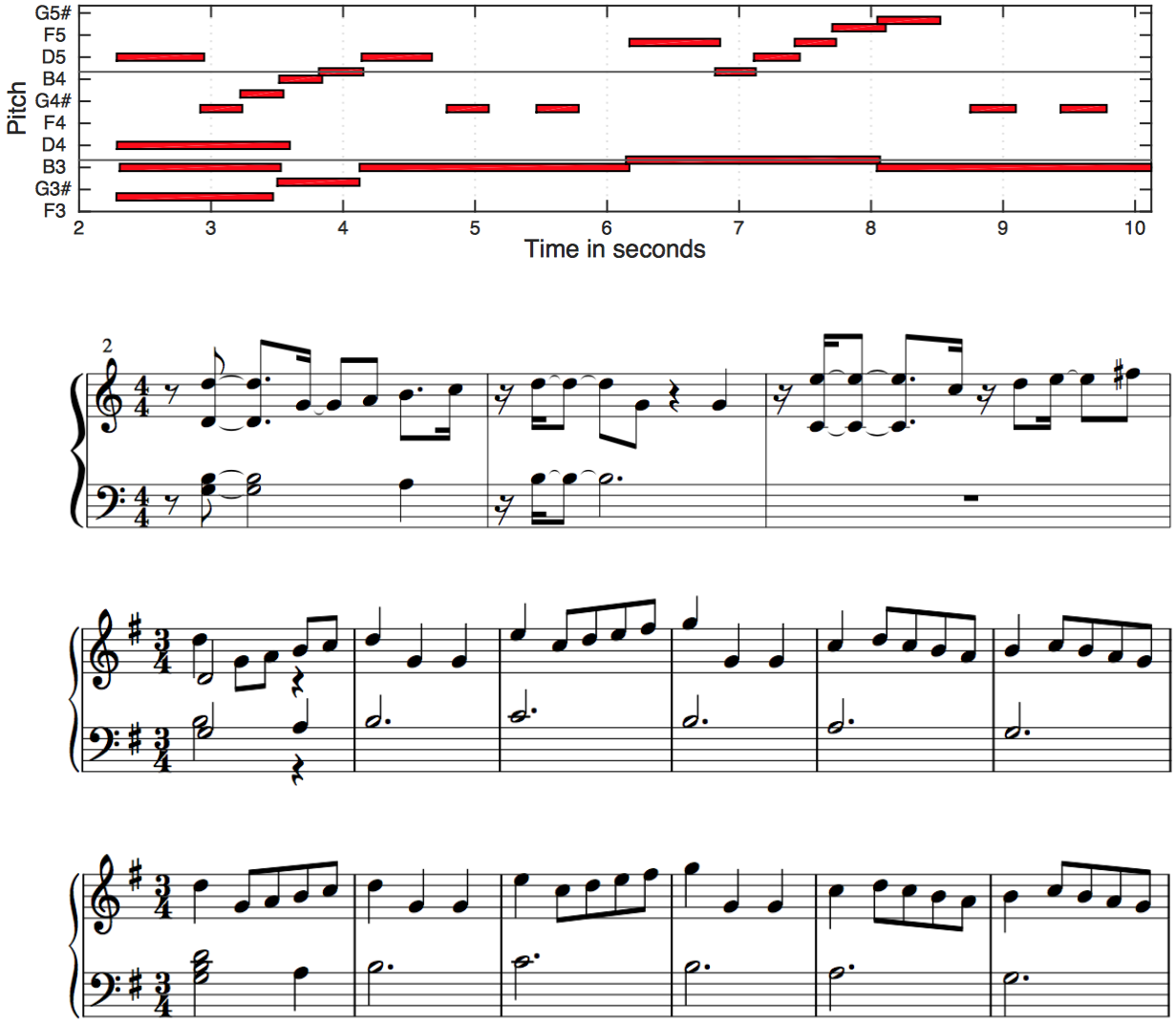
Also, some systems are able to:

- Detect & integrate rhythmic information
- Detect tuning (per piece/note)
- Extract velocity per detected note
- Transcribe fingering (for specific instruments)
- Quantise pitches over time/beats

Significant work needs to be done in order to extract a complete score

Tasks

Automatic Music Transcription



Tasks

Automatic Music Transcription

Factors:

- Notes:

  Spelling, Staff assignment, Group into chords

- Rests:

  Duration, Staff assignment

- Binary matching:

  Barlines, Clefs, Key signatures, Time signatures

Tasks

- Automatic Music Transcription
- **Music Synchronization**
- Source Separation
- Performance Expressiveness Analysis

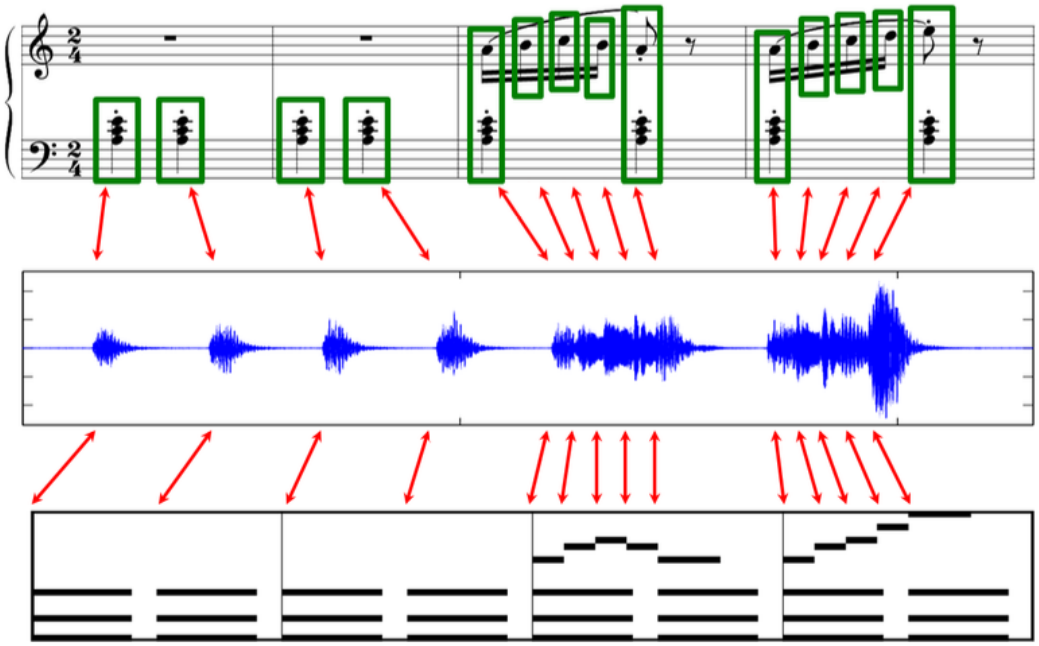
Tasks

Music Synchronization

Concept

Align different versions/modalities of music performance

- Different renderings of audio performances
- Music score (Sheet music)
- Video frames



Tasks

Music Synchronization

Categories

- Offline music alignment
  - Have the full sequence of both signals
  - Method: Dynamic Time Warping
- Realtime music alignment (score following)
  - Have the full sequence of one signal (music score)
  - The other signal comes as live streams
  - The system should find the alignment in real-time

Tasks

Music Synchronization

Applications of offline music alignment

- Query by Humming System
- Construct multi-modal music digital library

  https://jdasam.github.io/PerformScore/
- Music tutoring and grading

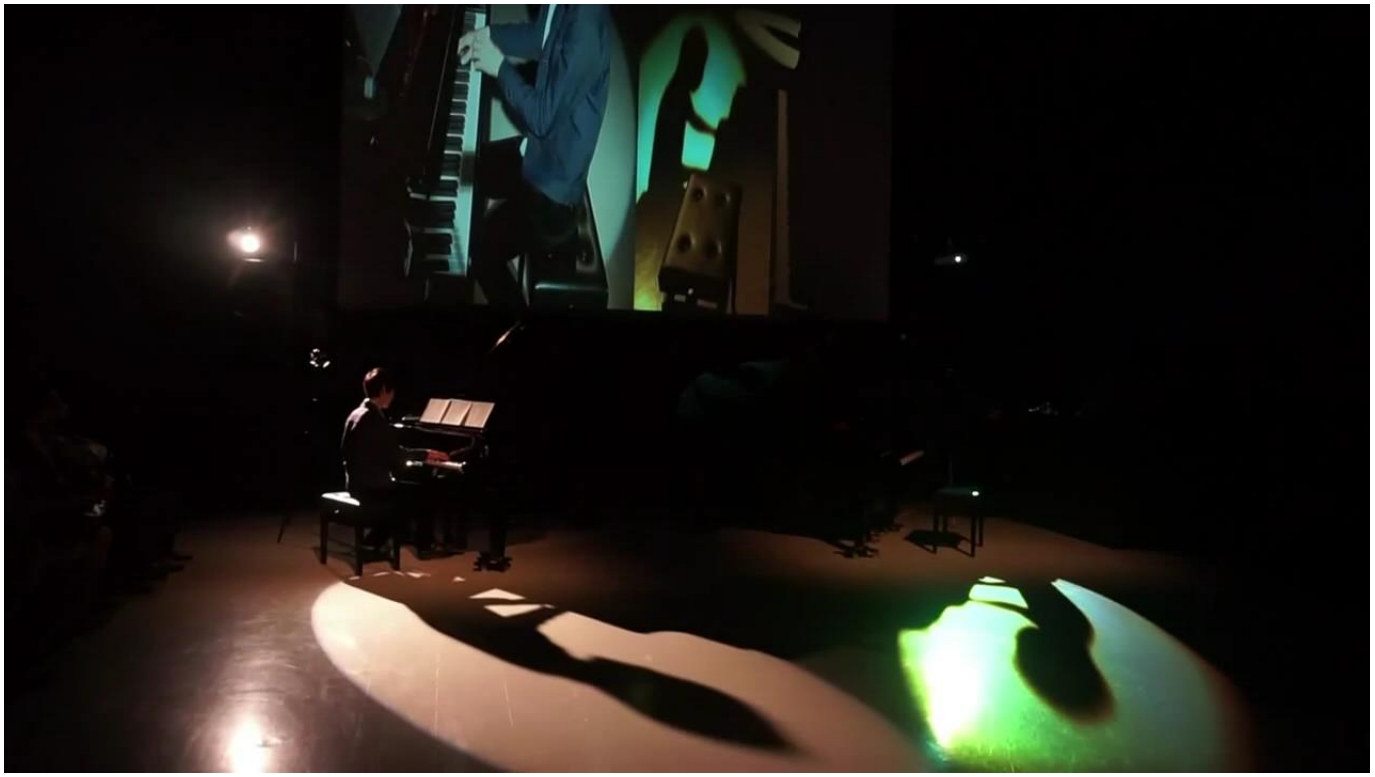
Tasks

Music Synchronization

Applications of real-time music alignment

- Automatic Accompaniment System

  - the computer follows the musician's speed



Tasks

Music Synchronization

Applications of real-time music alignment

- Automatic Accompaniment System

  - the computer follows the musician's speed

  https://www.youtube.com/watch?v=dnyJkwLzXPM

Tasks

Music Synchronization

Applications of real-time music alignment

- Automatic Lyrics Display

  - the computer follows the performance the display the encoded lyrics at the correct timing

Tasks

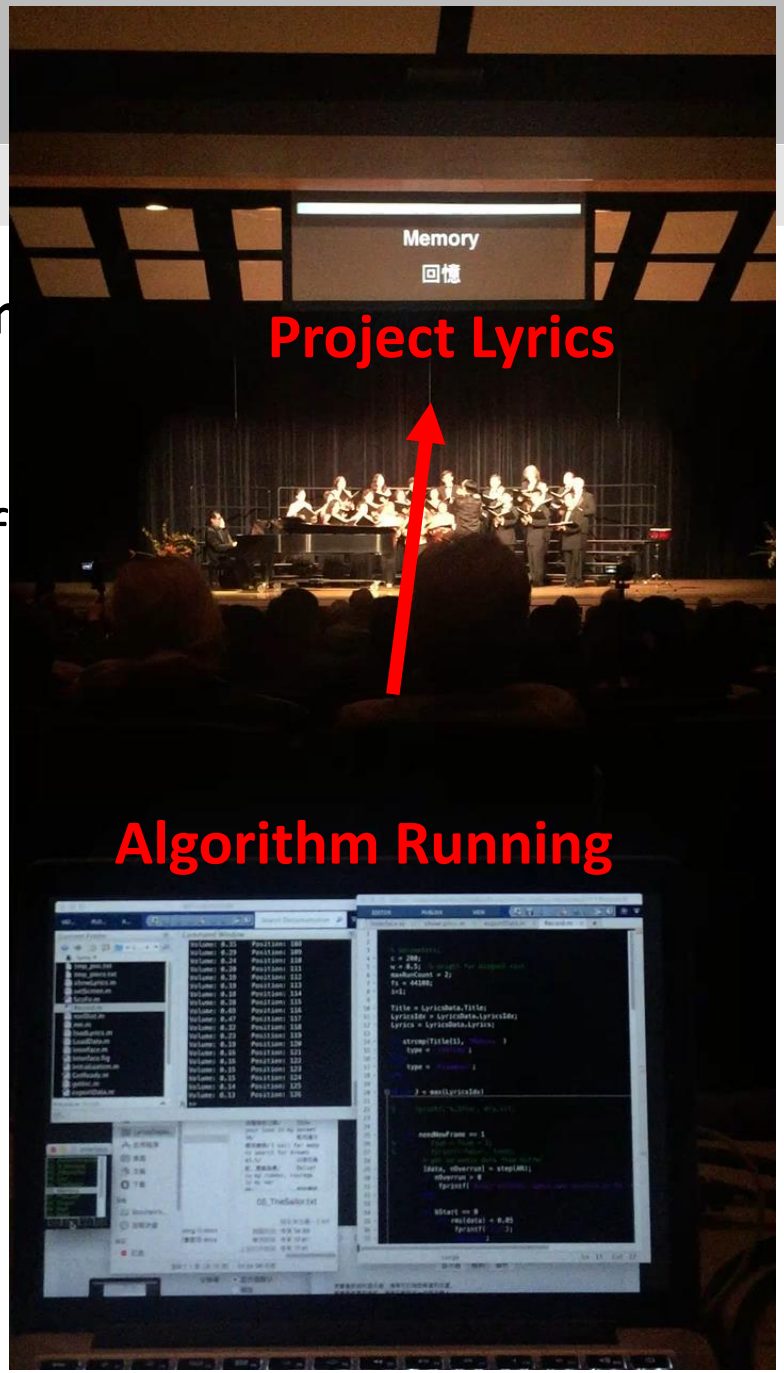
Music Synchronization

Applications of real-time music alignm

- Automatic Lyrics Display

  - the computer follows the perf

encoded lyrics at the correct timing



Tasks

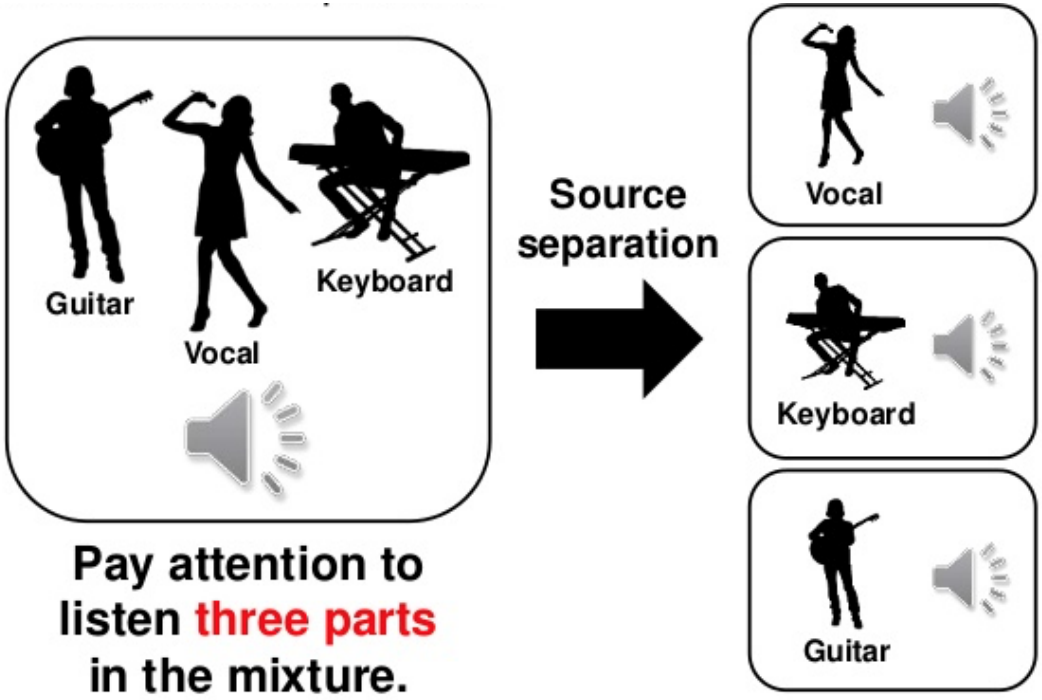
- Automatic Music Transcription
- Music Synchronization
- **Source Separation**
- Performance Expressiveness Analysis

Tasks

Source Separation

Concept

Separate the sound mixture into individual sources

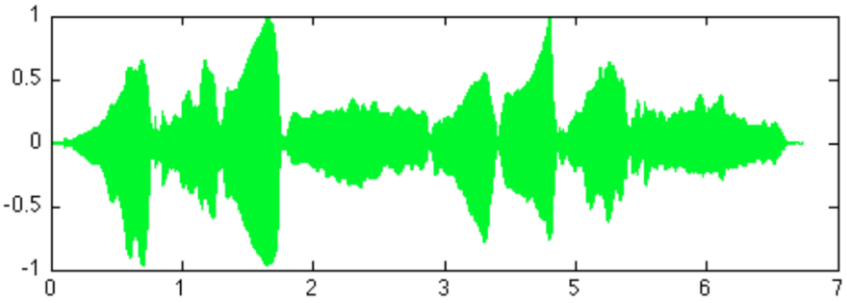


Tasks

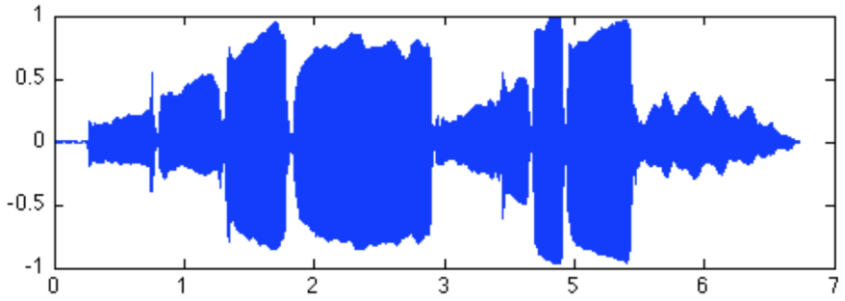
Source Separation

Time Domain

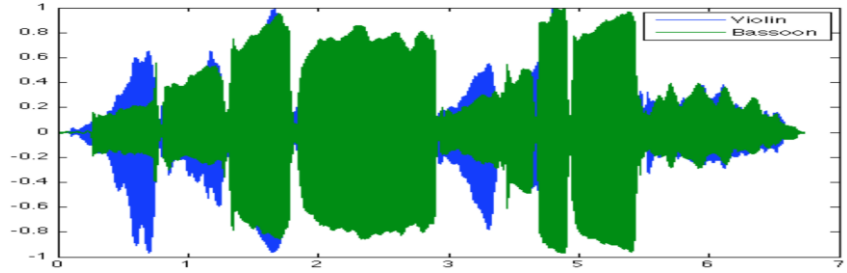
Violin

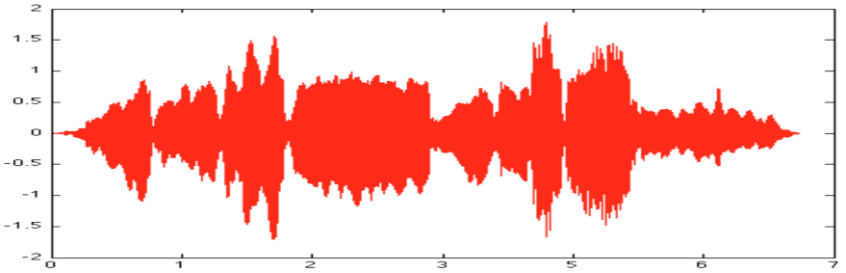


Bassoon



Mixture





It is not easy!

Tasks

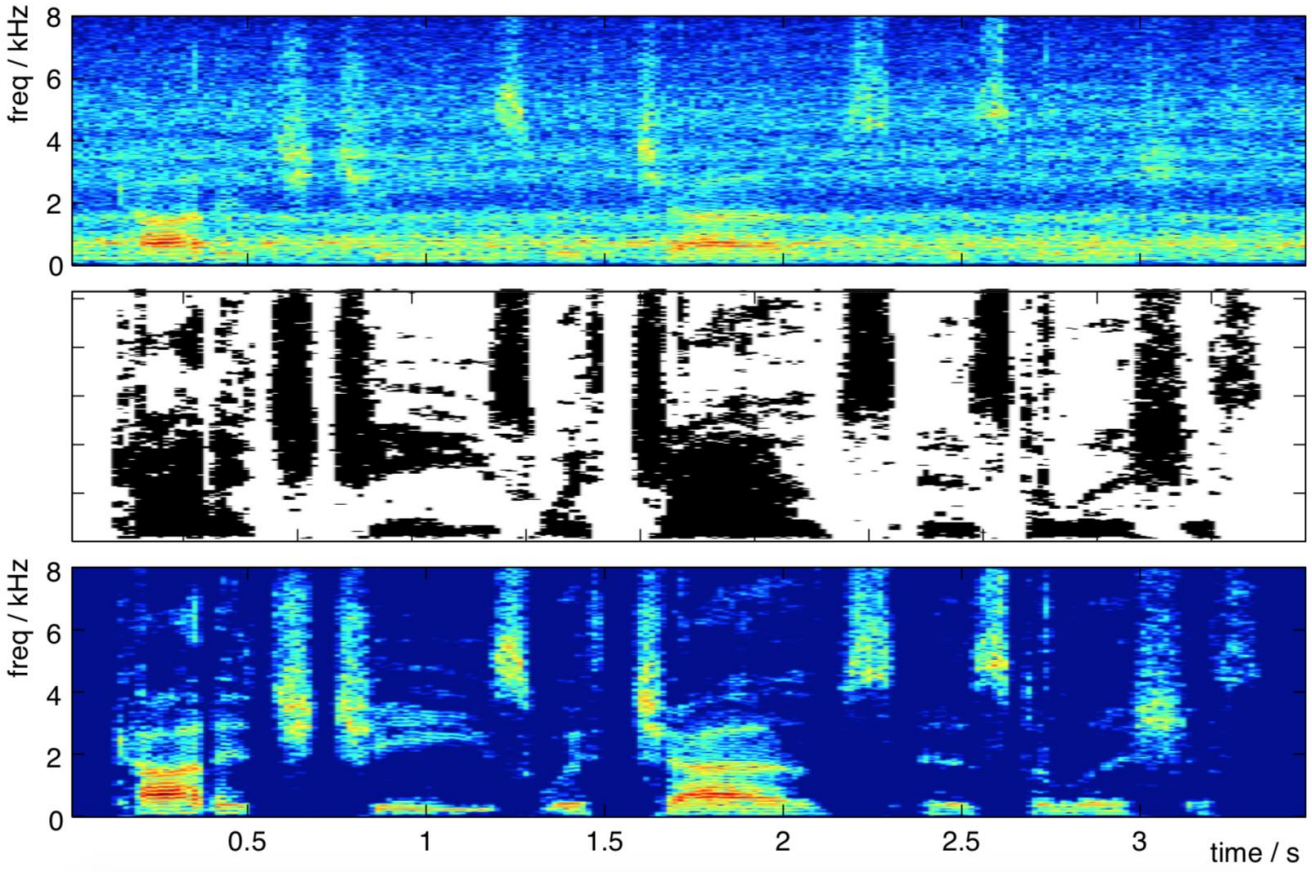
Source Separation

Frequency Domain

Spectrogram of sound mixture

Mask

Spectrogram of separated source



The trick is to find the right mask

Tasks

Source Separation

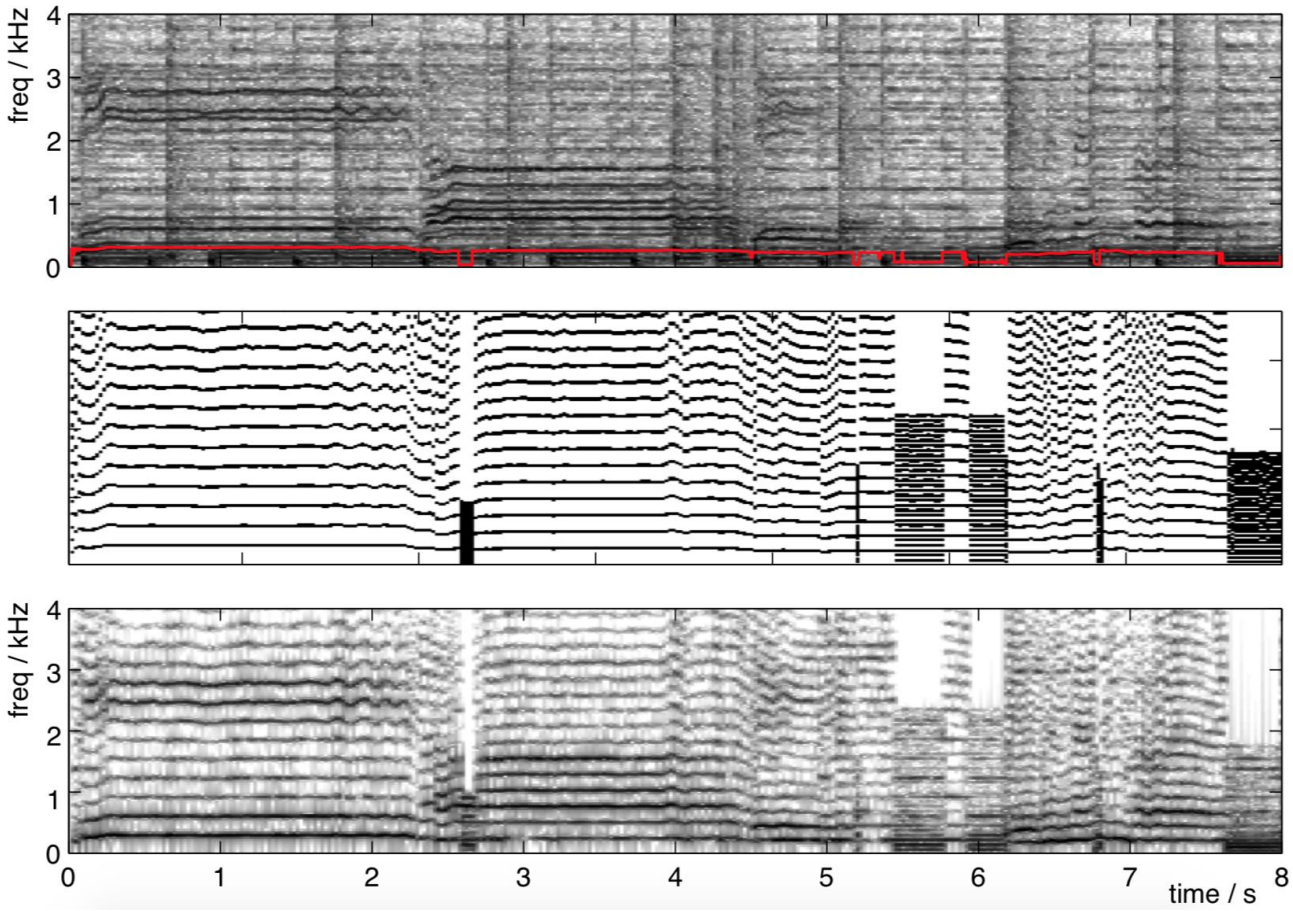
Harmonic mask:

Given the pitch track, know where to expect harmonics

Spectrogram of sound mixture

Mask

Spectrogram of separated source



Tasks

Source Separation

Results:

- https://www.youtube.com/watch?v=b07t-y1jNcs
- https://www.youtube.com/watch?v=C8aZYVaFcjE
- http://paris.cs.illinois.edu/demos/ai/user-guide.mp4

Tasks

- Automatic Music Transcription
- Music Synchronization
- Source Separation
- **Performance Expressiveness Analysis**

Tasks

Performance Expressiveness Analysis

What is expressiveness?

- Volume
- Tempo
- Legato/Staccato/Vibrato
- Up-bow/Down-bow (String instrumentalists)
- Body Movements

Some are visual aspect of music performance

Tasks

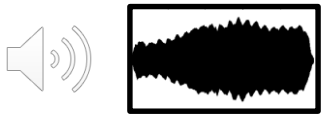
Performance Expressiveness Analysis

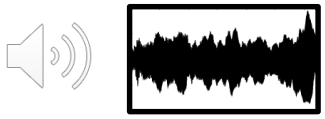
Vibrato Analysis

- Important artistic effect
- Pitch modulation of a note in a periodic fashion
- Characterized by Rate & Extent

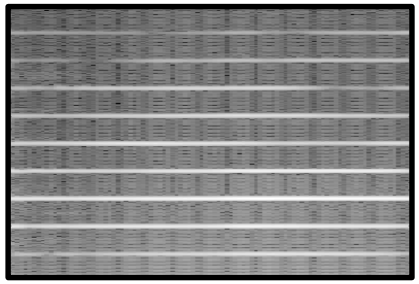


Audio

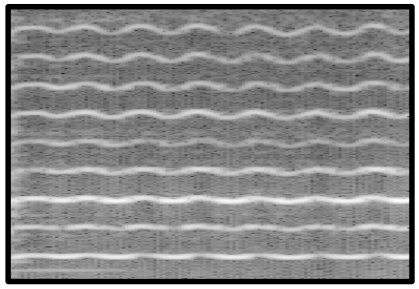




Spectrogram



Non-vibrato



Vibrato

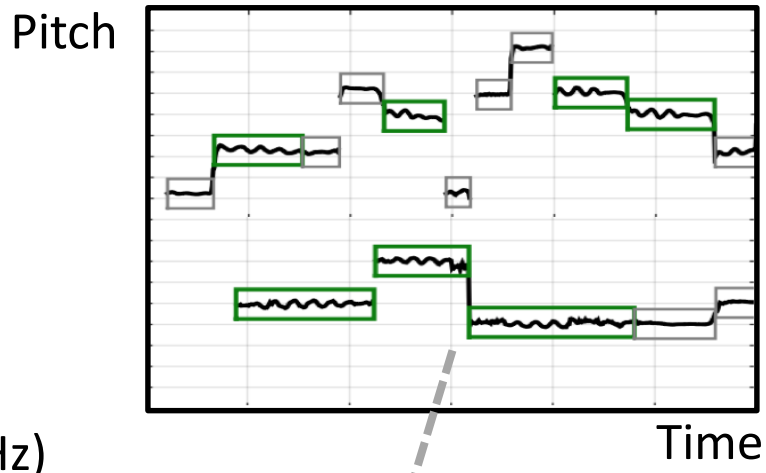
Tasks

Performance Expressiveness Analysis

Vibrato Analysis

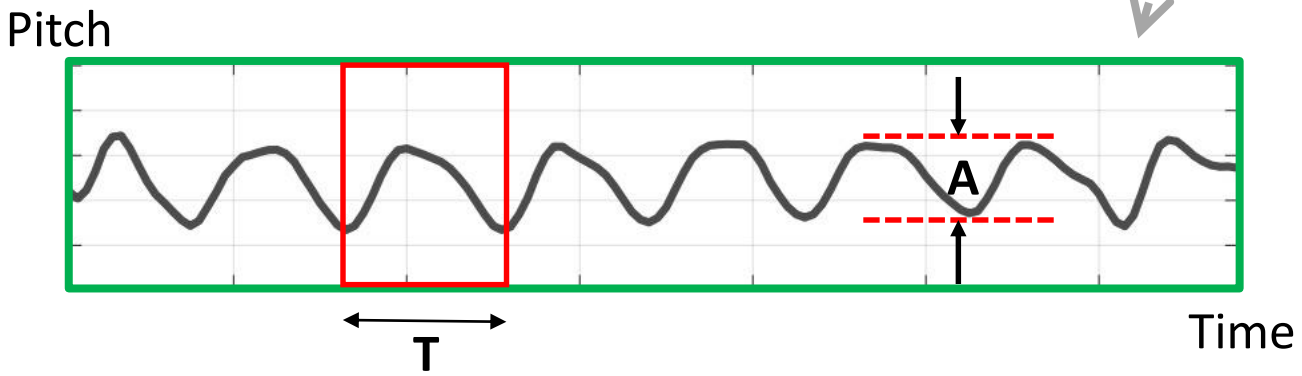
Vibrato Detection

- Note-level vibrato/non-vibrato classification



Vibrato Characterization

- Vibrato rate: speed of pitch variation (1/T Hz)
- Vibrato extent: amount of pitch variation (A cents)



Tasks

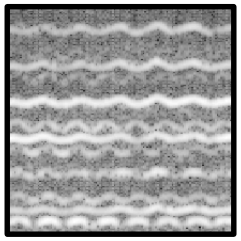
Performance Expressiveness Analysis

Vibrato Analysis from visual modality

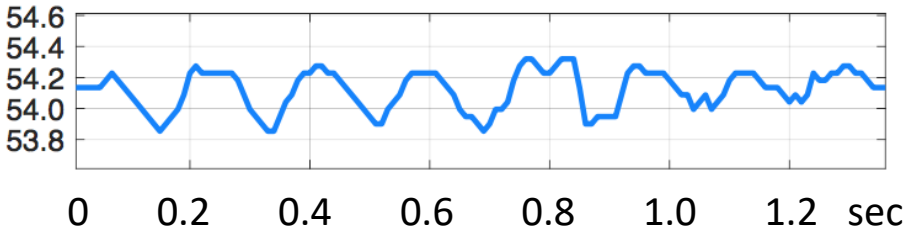
Audio-based, Polyphonic



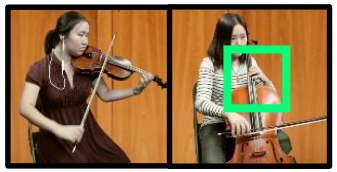
Spec



Pitch



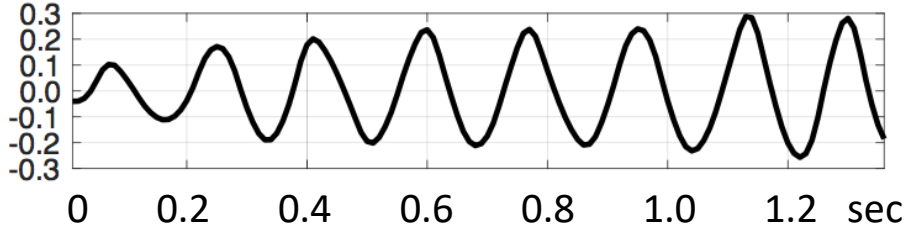
Video-based



Hand



Hand Displacement



Tasks

Performance Expressiveness Analysis

Visually Onset Prediction

- Predict the onset from body motion
- Help a robot accompanying player better synchronize with human



Tasks

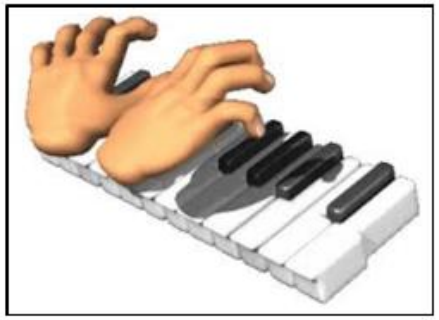
Performance Expressiveness Analysis

Body Expressiveness Modeling and Generation

Applications

- Visual expressions give immersive music enjoyment experiences
- Replicating musicians' body motions for educational purposes
- Visual human-computer interactions in automatic accompaniment system





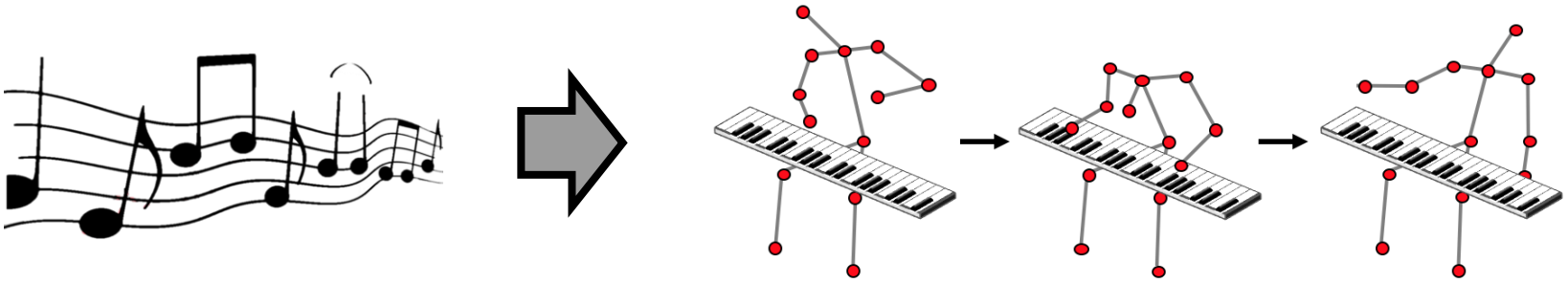


Tasks

Performance Expressiveness Analysis

System 1: MIDI to Pose

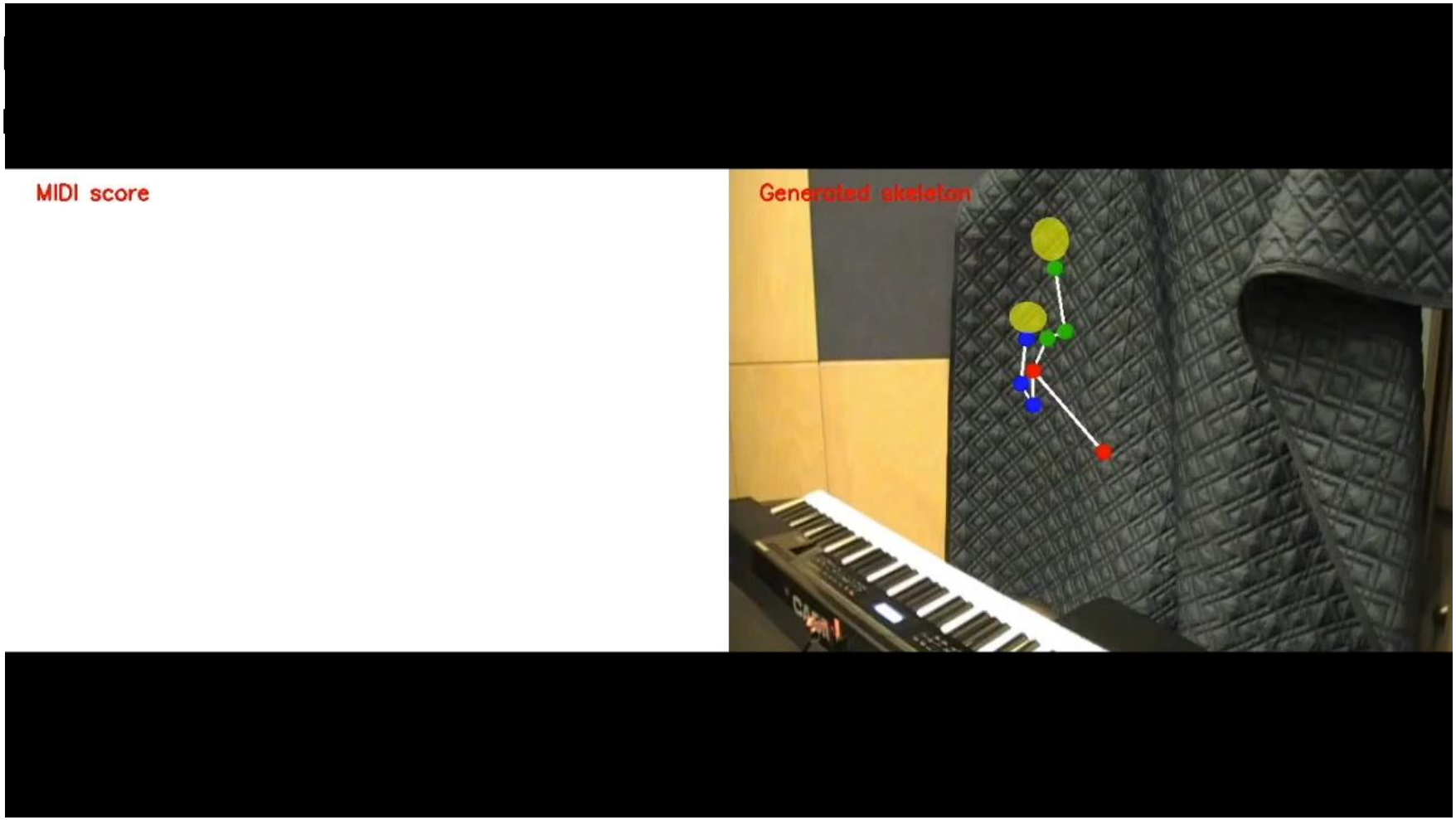
Input a sheet music, and output a simulation of human's body motion of playing this piece



Tasks

Performance Expressiveness Analysis

System 1: MIDI to Pose

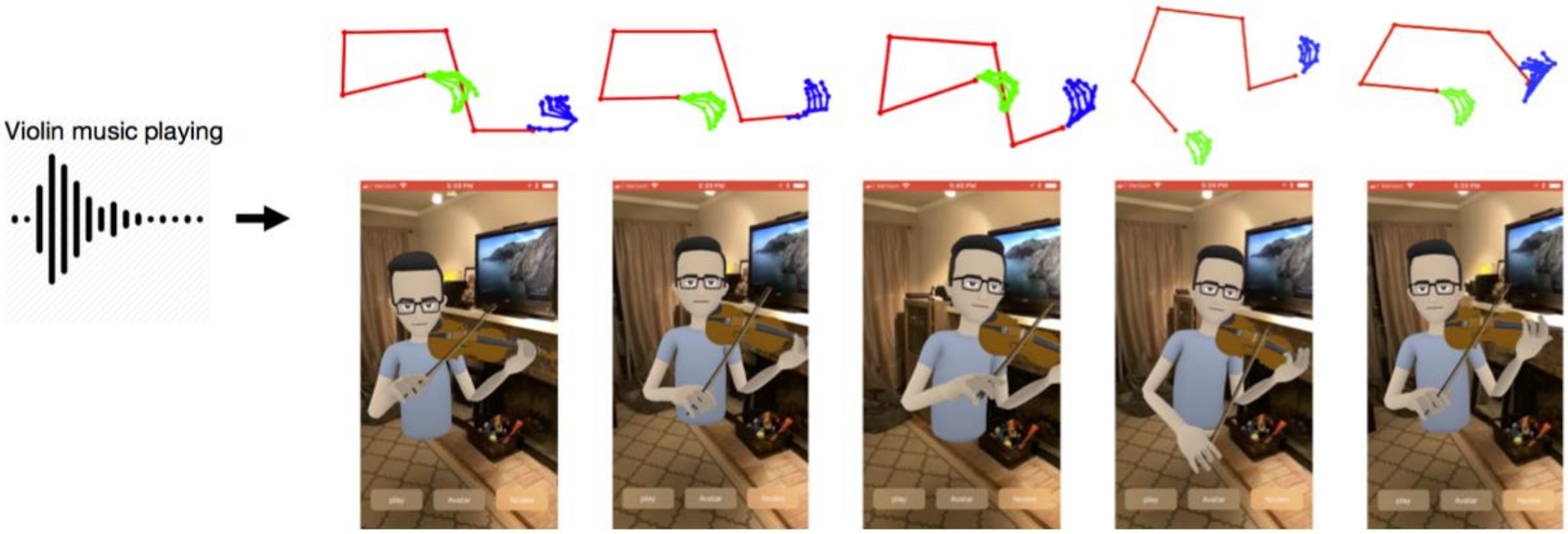


Tasks

Performance Expressiveness Analysis

System 2: Audio to Body Dynamics

Input an audio performance, and output a simulation of human's body motion of playing this piece



MIR Community

- International Society for Music Information Retrieval (ISMIR)
- The International Conference on New Interfaces for Musical Expression (NIME)
- Sound and Music Computing (SMC)
- International Computer Music Conference (ICMC)
- MIR Evaluation Exchange (MIREX)
  http://www.music-ir.org/mirex/wiki/MIREX_HOME
- MIR-related PhD thesis
  http://www.pampalk.at/mir-phds/