UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat d'Informàtica de Barcelona

FIB

Wave2Page

Gerard Caravaca, Armando Rodriguez and Benjami Parellada

Midterm Presentation - Intelligent System Project - 2023-2024

Master in Artificial Intelligence (UPC-UB-URV)

Background

Key Concepts:

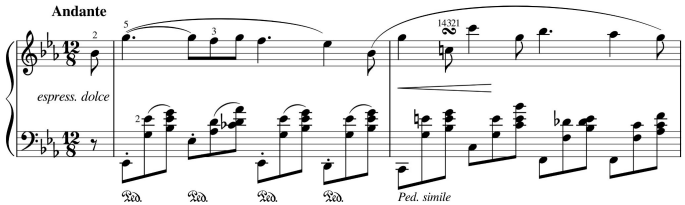
- **MIDI (Musical Instrument Digital Interface):** Protocol that describes electronic instruments to connect and communicate with each other
- **Music Transcription:** Song performance to sheet music
- **Automatic Music Transcription:** Use of tech to convert music into sheet music

Importance of AMT:

- **Education:** Facilitates learning music pieces
- **Preservation:** Converts performances to an easily accessible format
- **Creative Aid:** Assist composers into experimenting

Transcribe

Performance



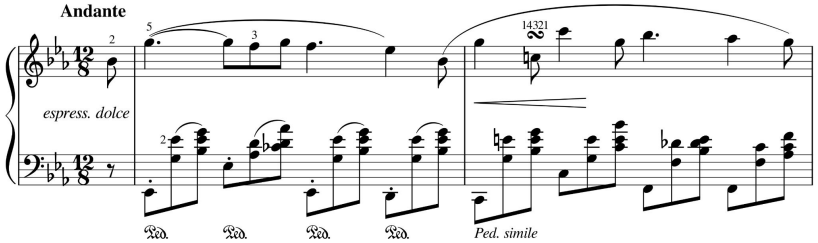
Context

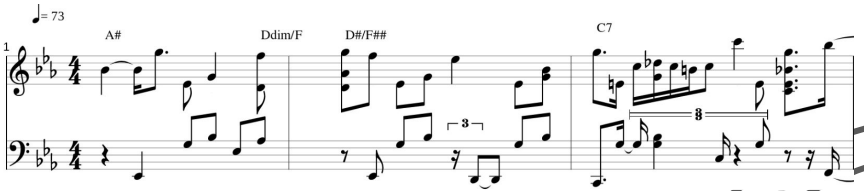
Project Focus:

- Create IS capable of converting audio recordings into sheet music.
- **Integrating disciplines:** Signal Processing, Deep Learning, and Music Theory

Problem Definition:

- Transcribing music is **complex and intricate**
- Current automatic systems **rely on MIDI** (lack finesse of human transcriptions)
- **Challenges:**
  - **Over-fidelity in MIDI** causing overly complex transcriptions.
  - Audio to MIDI conversion introduces **artifacts**





Previous Solutions

AMT types:

- Old discipline spanning 40 years
- Many various types and solutions proposed: Fourier, Neural Networks, etc.

AMT depth:

- Different levels of abstraction: Frame, Note, Stream
- Each has its own strengths and weaknesses

Market Analysis:

- ScoreCloud, MuseScore, AnthemScore, NoteFlight
- Digital Audio Workstation: Cubase
- Still prone to errors, not end to end

Main Contribution

Main Contribution:

- Emphasized End-to-End system
- Use State-of-the-Art, see if we can advance on any front

Advantages Over Existing Systems:

- Poly instrument sheet generation
- Curated sheet music, reminiscent of Human
- Drum and Vocal transcription

Potential Impact:

- Creation of a pipeline that can proliferate multiple different tools and applications

Requirements

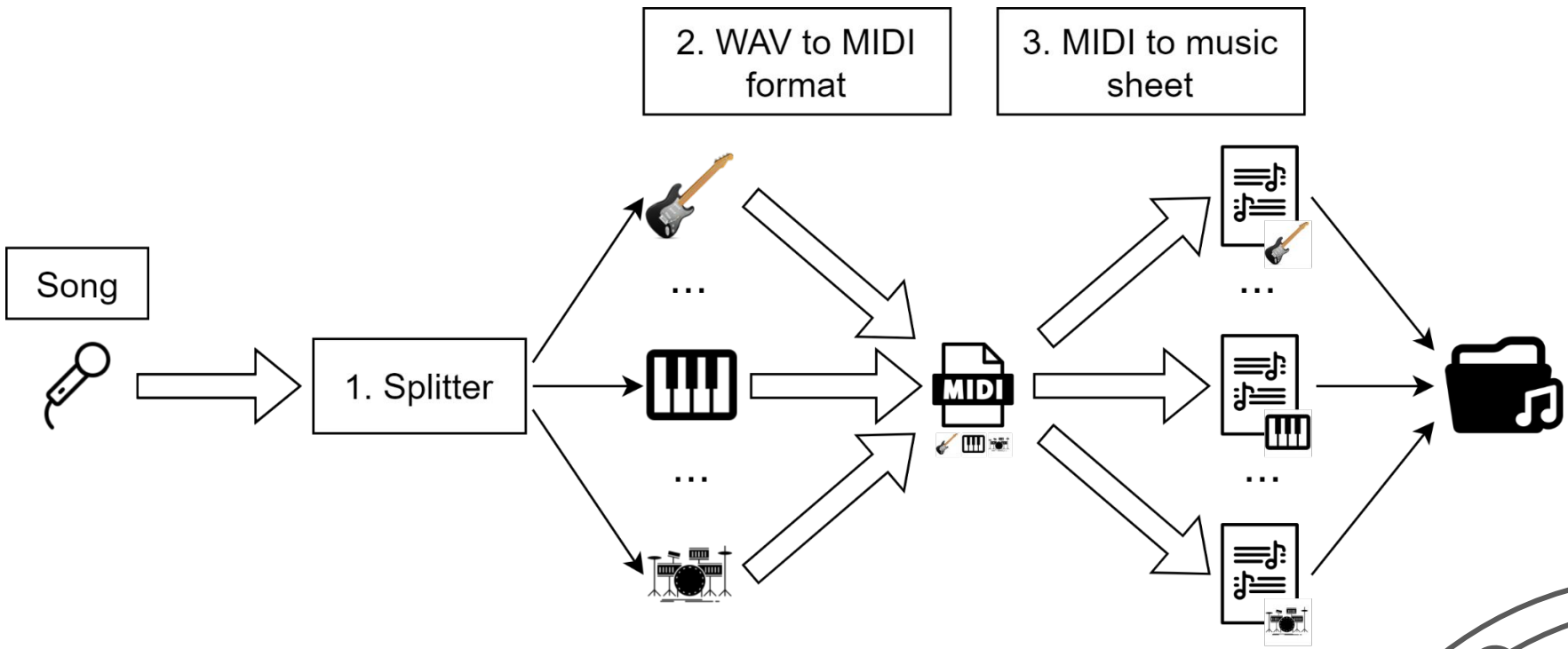
Functional

- Multiple Instruments and tracks
- Machine Learning and AI
- Format compatibility
- Feedback mechanism
- Documentation and support

Non-functional

- Accuracy
- Efficiency
- Robustness
- Scalability
- Accessibility
- ...

Solution design



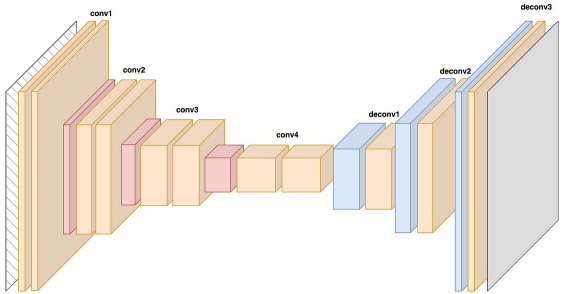
Splitter

Music Source Separation (Demucs)

Specifically, the htdemucs_6s model by facebookresearch

Advantages

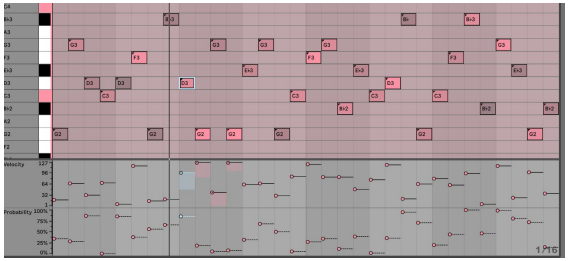
- State-of-the-art
- 6 sources separation
- Allows gpu acceleration



Limitations

- Piano separation is not the best and could be improved
- Background noise can be detected in some samples
- 1.5 times the duration of the track using cpu

Wave2Midi



Automatic Music Transcription (Basic Pitch)

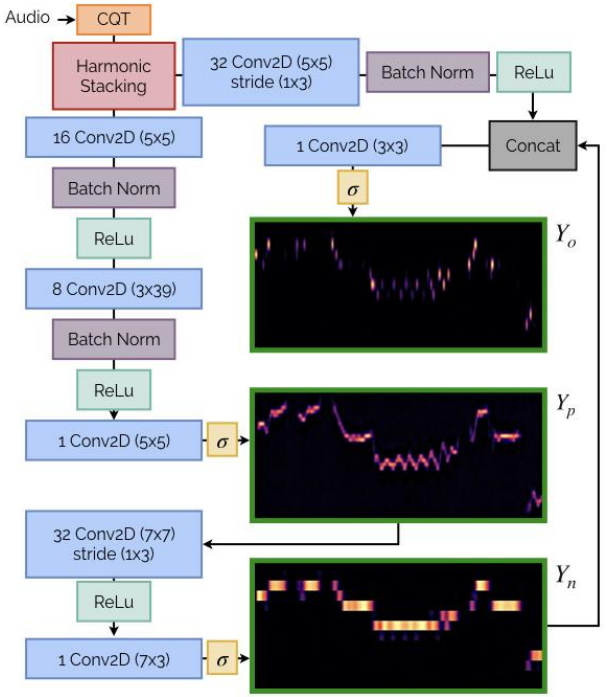
Developed by Spotify

Advantages

- State-of-the-art
- Instrument-Agnostic
- Allows gpu acceleration
- Lightweight

Limitations

- Note creation event is based on heuristics
- Worse in piano and voice



Midi2Sheet

MuseScore4 as the core converter:

- Supports a wide range of musical notations
- Extensive set of features to achieve our requirements
- Generation quality, robustness and flexibility
- From simple compositions to complex scores
- Grouped and individual music sheets

Risk analysis

Risk	Likelihood	Impact	Severity
Time management	POSSIBLE	TOLERABLE	MEDIUM
Unforeseen results	PROBABLE	UNDESIRABLE	HIGH
Hardware limitation	POSSIBLE	TOLERABLE	MEDIUM
Design failures	POSSIBLE	INTOLERABLE	EXTREME
Ethical concerns	IMPROBABLE	UNDESIRABLE	MEDIUM
Legal issues	POSSIBLE	INTOLERABLE	EXTREME
Financial issues	IMPROBABLE	ACCEPTABLE	LOW
New technologies arise	PROBABLE	ACCEPTABLE	MEDIUM

Next steps

- **Splitter post-processing**
  - Piano separation dedicated DL model (vocal remover)
  - Background noise removal - Spectral subtraction
- **Wav2Midi improvement**
  - Delete sound artifacts by post-processing the Midi file
- **Midi2Sheet customization**
  - Improve the sheets' visualization to make them more attractive
- **Web page creation**
  - Create a user-friendly Flask web page that englobes all the steps

Demo: Splitter

Original audio

Divided audio

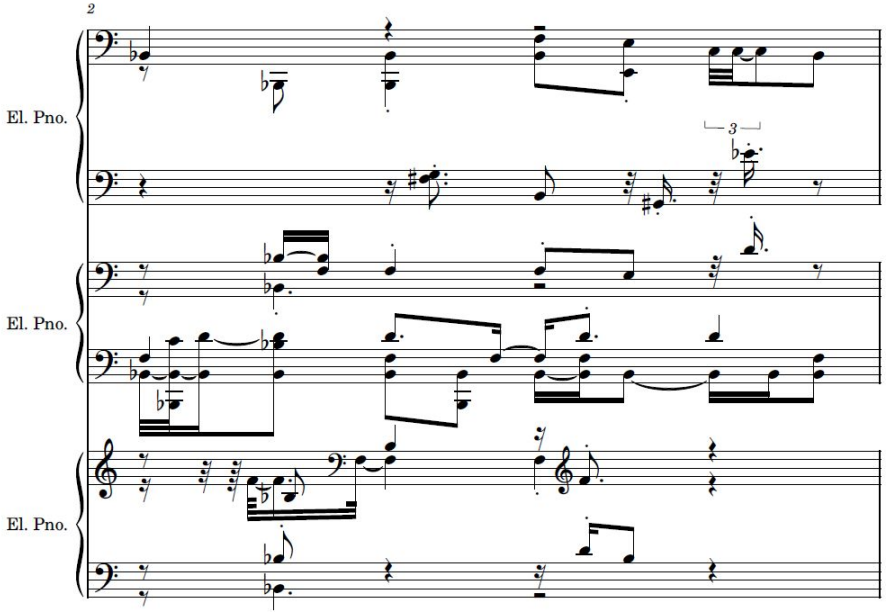
Bass

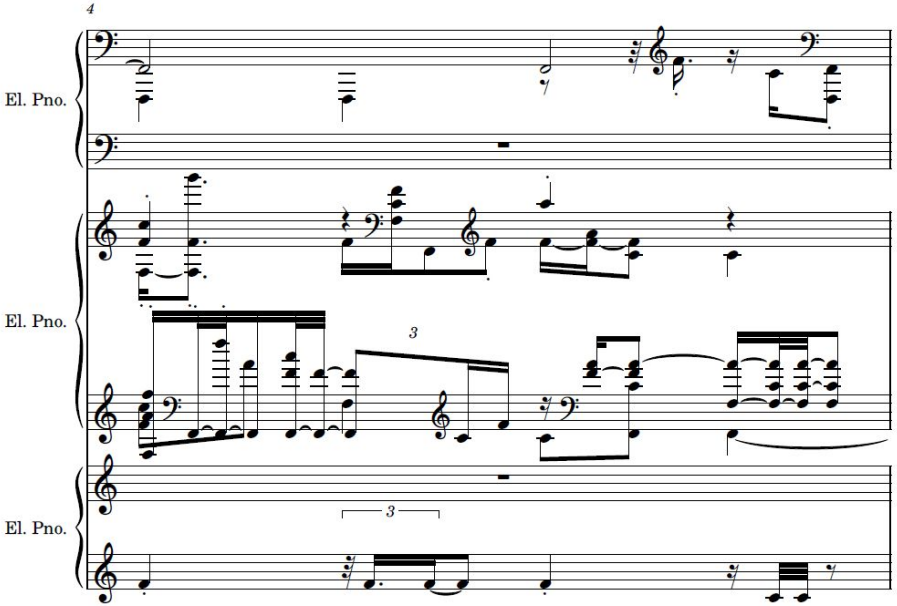
Drums

Piano

Vocals

Demo: Music sheets





UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat d'Informàtica de Barcelona

FIB

Wave2Page

Gerard Caravaca, Armando Rodriguez and Benjami Parellada

Midterm Presentation - Intelligent System Project - 2023-2024

Master in Artificial Intelligence (UPC-UB-URV)