Towards Automatic Instrumentation by Learning to Separate Parts in Multitrack Music

Hao-Wen Dong[1] Chris Donahue[2] Taylor Berg-Kirkpatrick[1] Julian McAuley[1]

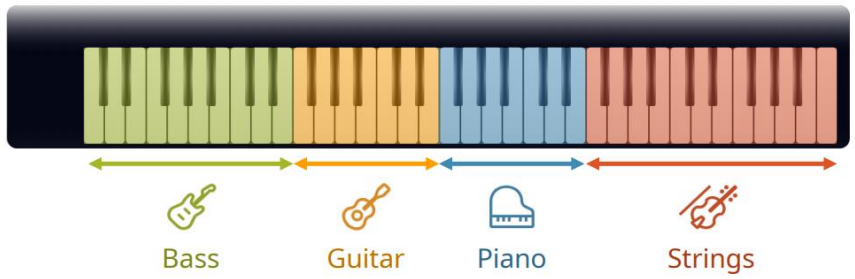
[1] University of California San Diego [2] Stanford University



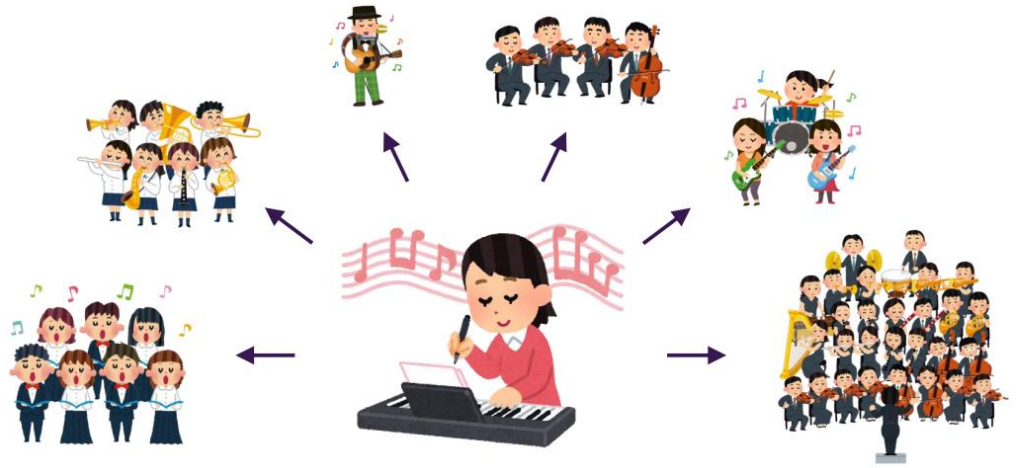

UC San Diego Stanford

Automatic instrumentation

- **Goal**—Dynamically assign instruments to notes in solo music



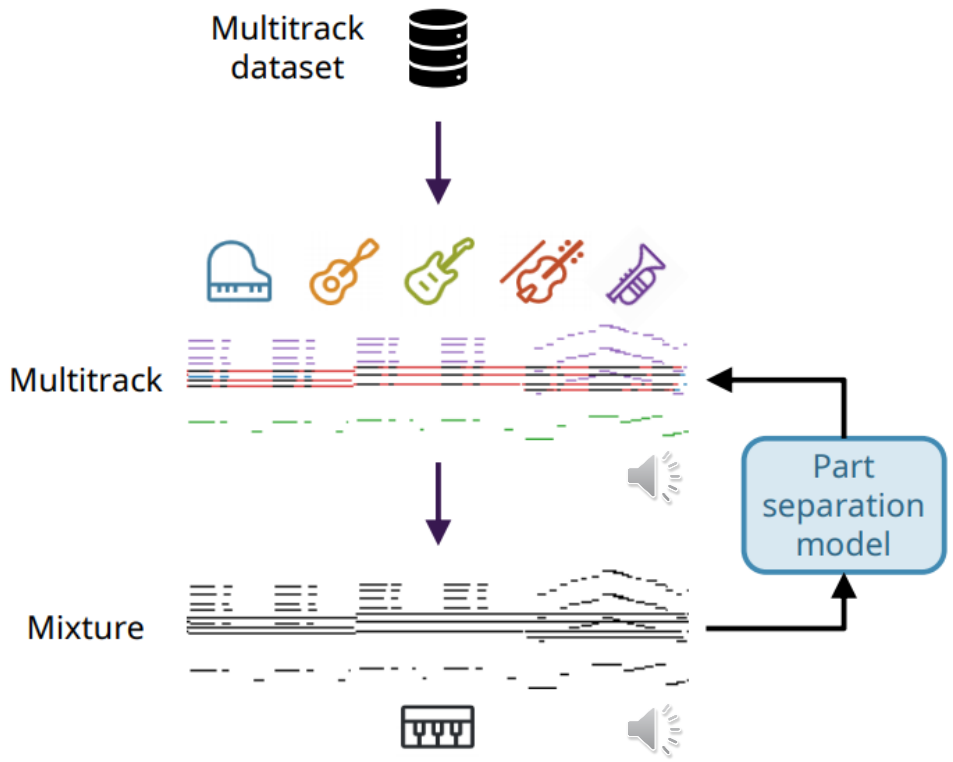
Intelligent musical instruments



Assistive composing tools

Overview

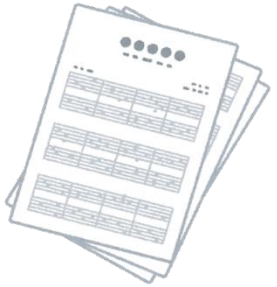
- Acquire **paired data**
  - Downmix multitracks into single-track mixtures
- Train a **part separation** model
  - Learn to infer the part label for each note in a mixture
- Approach **automatic instrumentation**
  - Treat the input as a mixture
  - Separate out the relevant parts



Data

- Four datasets of diverse genres and ensembles

Dataset	Hours	Files	Notes	Parts	Ensemble	Most common label
Bach chorales [31]	3.23	409	96.6K	4	soprano, alto, tenor, bass	bass (27.05%)
String quartets [32]	6.31	57	226K	4	first violin, second violin, viola, cello	first violin (38.72%)
Game music [33]	45.05	4.61K	2.46M	3	pulse wave I, pulse wave II, triangle wave	pulse wave II (39.35%)
Pop music [34]	1.02K	16.2K	63.6M	5	piano, guitar, bass, strings, brass	guitar (42.50%)



Data representation

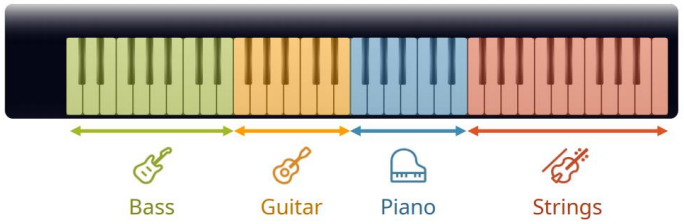
A **sequence of notes** specified by

- **Time**—onset time (in time step)
- **Pitch**—pitch as a MIDI note number
- **Duration**—note length (in time step)
- **Frequency**—frequency of the pitch (in Hz)
- **Beat**—onset time (in beat)
- **Position**—position within a beat (in time step)

Model

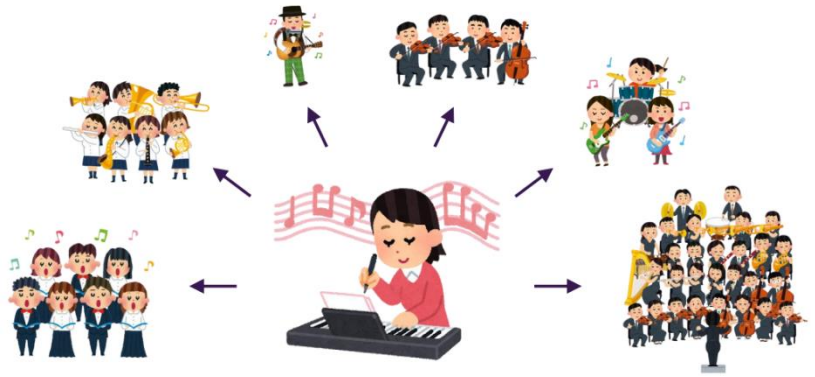
Online models

- LSTMs
- Transformer decoders



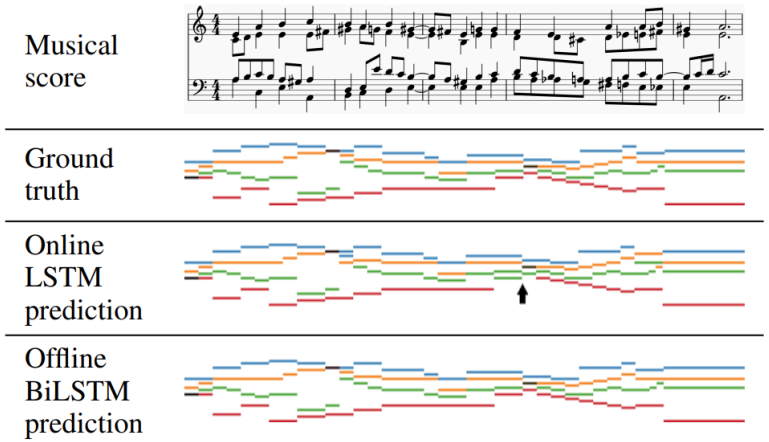
Offline models

- BiLSTMs
- Transformer encoders



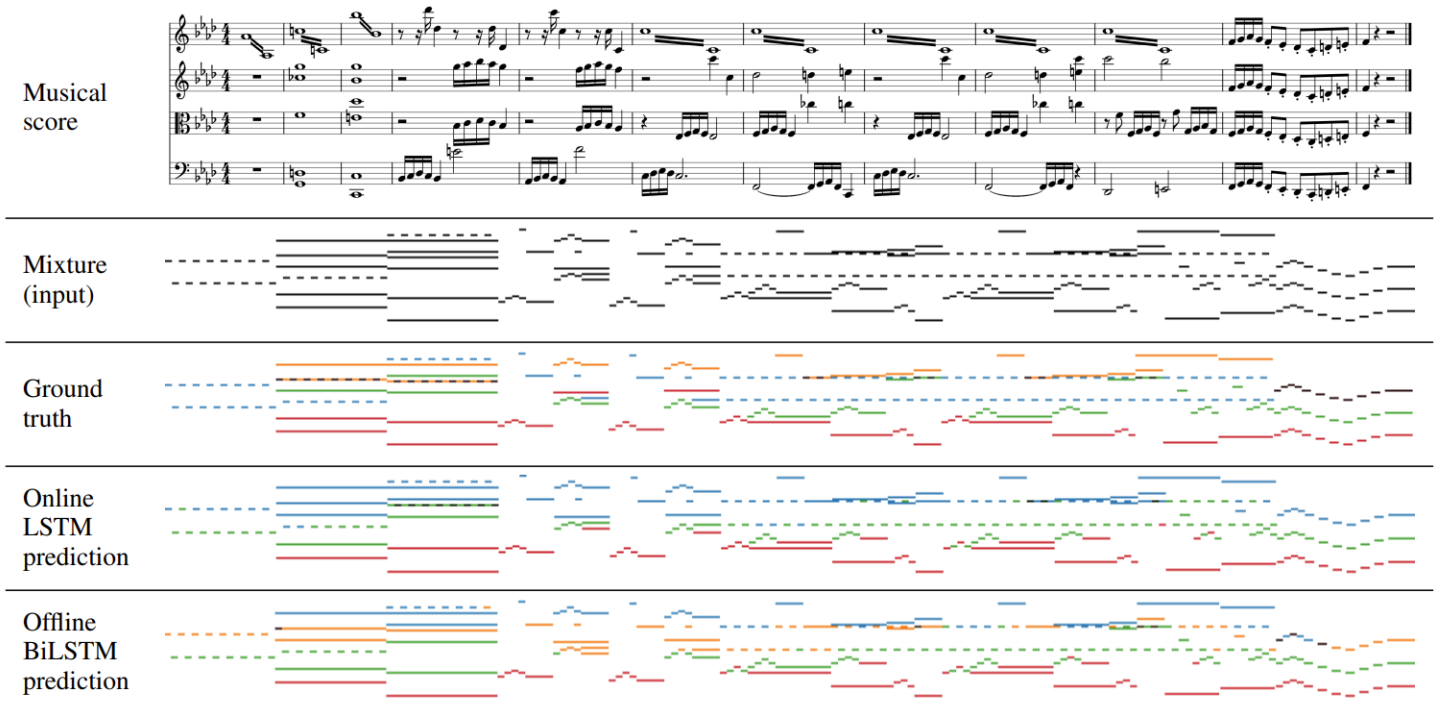
Results

Bach chorales



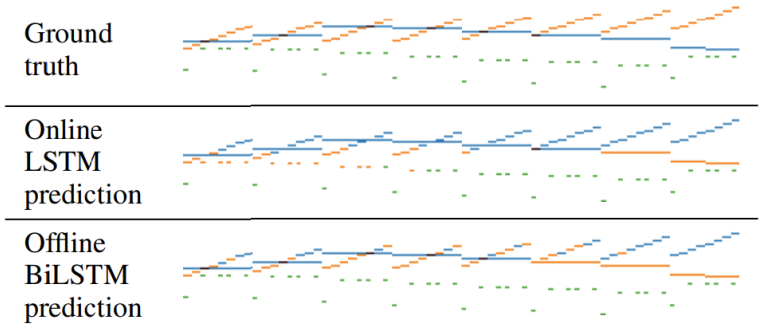
(Audio available. [1] Colors: soprano, alto, tenor, bass.)

String quartets



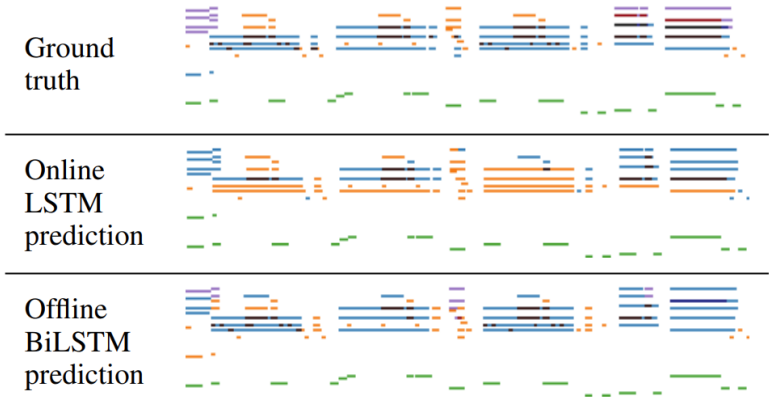
(Audio available. [1] Colors: first violin, second violin, viola, cello.)

Game music



(Audio available. [1] Colors: pulse wave I, pulse wave II, triangle wave.)

Pop music



(Audio available. [1] Colors: piano, guitar, bass, strings, brass.)

Demo

- Our proposed models can produce alternative convincing instrumentations for an existing arrangement.

Original

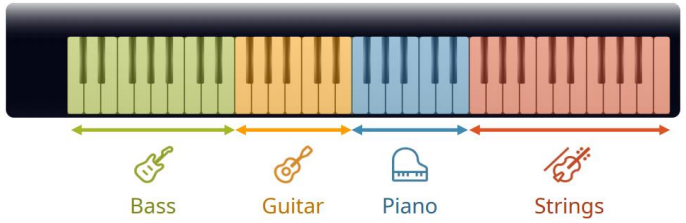
LSTM without entry hints

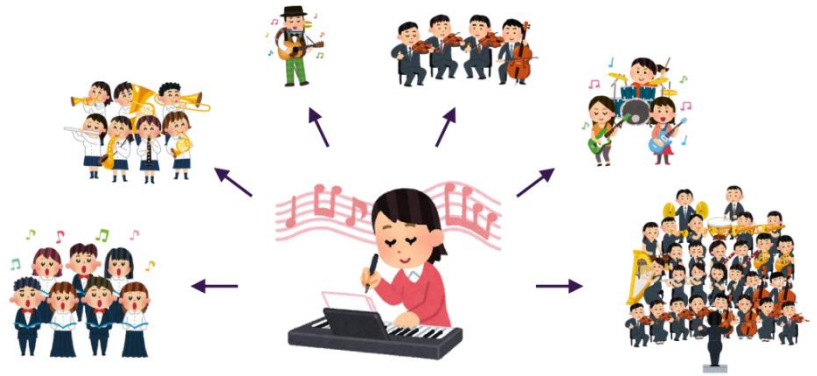
BiLSTM with entry hints

(Audio available. [1] Colors: piano, guitar, bass, strings, brass.)

Summary

- Proposed a new task of part separation
- Showed that our proposed models outperform various baselines
- Presented promising results for applying a part separation model to automatic instrumentation





Thank you!

Learn more at https://salu133445.github.io/arranger/