

Timbre Comparison in Note Tracking from Onset, Frames and Pitch Estimation

Carlos Hernandez-Olivan, Ignacio Zay Pinilla, Jose R. Beltran

Instituto de Investigación en Ingeniería de Aragón (I3A)
 Universidad de Zaragoza, Mariano Esquillor s/n, 50018, Zaragoza, Spain.

Abstract

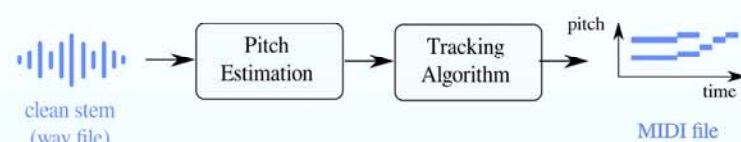
Note Tracking (NT) is a subtask of Automatic Music Transcription (AMT) which is a critical problem in the field of Music Information Retrieval (MIR). The aim of this work is to compare the performance of two models, one for onsets and frames prediction and another one with pitch detection and a note tracking algorithm in order to study the behaviour of different timbres and families of instruments in note tracking subtasks.

INTRODUCTION

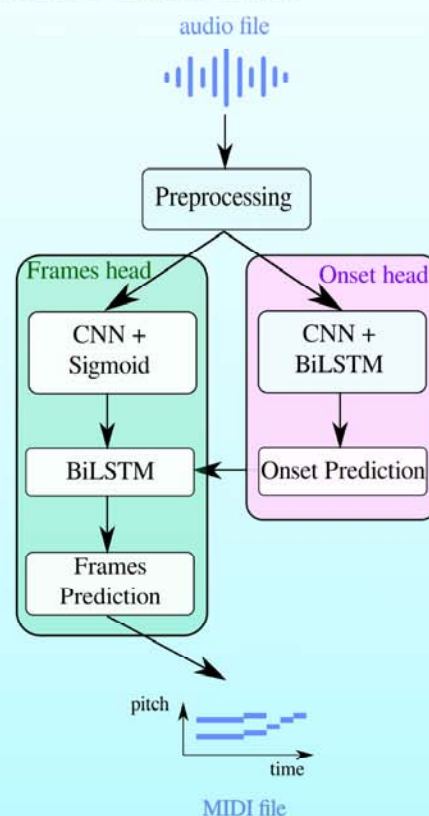
The Automatic Music Transcription problem can be separated into several subtasks, including multipitch estimation or frame-level transcription on pitches (MPE), note-level transcription on pitches, onset, and duration, also known as note tracking (NT) or instruments identification. Although transcribing a monophonic recording is considered to be a solved problem, ATM still remains an open research problem when it comes to multiple instruments (mixed signals) and polyphonic music.

METHODS

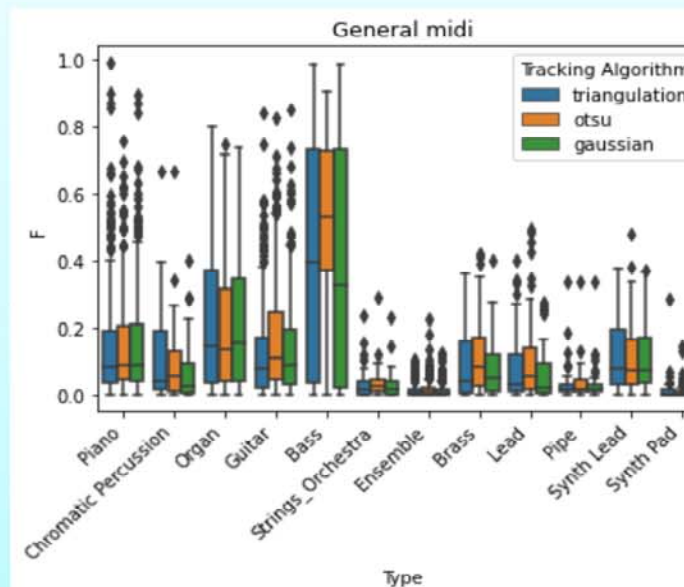
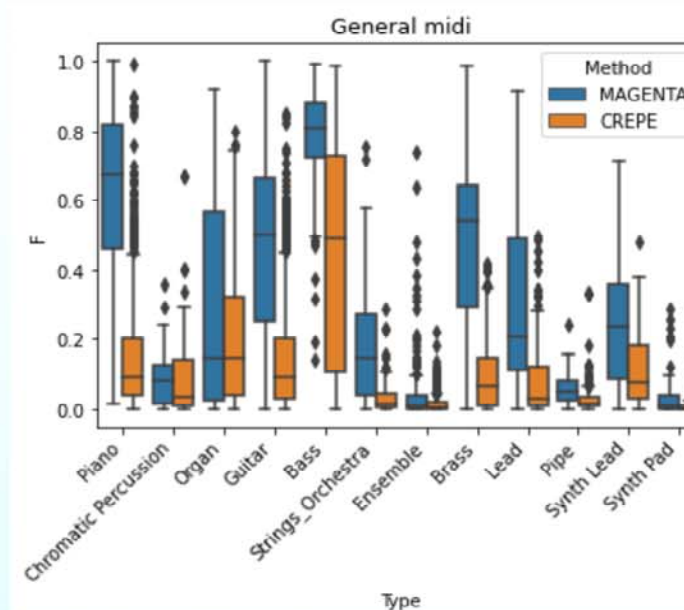
CREPE and Note Tracking Algorithm



Magenta Onsets and Frames



RESULTS



Instrument	Method	Note F ₁
Bass	Note Tracking (Otsu)	0.5575
	Magenta OaF	0.6694
Guitar	Note Tracking (Otsu)	0.3168
	Magenta OaF	0.6432
Synth pad	Note Tracking (Otsu)	0.0679
	Magenta OaF	0.1842
Synth lead	Note Tracking (Otsu)	0.3016
	Magenta OaF	0.3459
Brass	Note Tracking (Otsu)	0.3215
	Magenta OaF	0.5899
Strings	Note Tracking (Otsu)	0.1631
	Magenta OaF	0.4888
Organ	Note Tracking (Otsu)	0.2826
	Magenta OaF	0.2767
Piano	Note Tracking (Otsu)	0.3205
	Magenta OaF	0.9068
Chromatic Percussion	Note Tracking (Otsu)	0.2464
	Magenta OaF	0.5627

CONCLUSIONS

This work shows an overview of how different timbres affect some subtasks of automatic music transcription such as note tracking from an estimated f0 or onsets and frames prediction. By comparing the results, we can see that timbre and instrument onsets are variables that affect the results of music transcription in different subtasks.