AUDIO LABS

Modeling and Estimating Local Tempo: A Case Study on Chopin's Mazurkas

Hendrik Schreiber, Frank Zalkow, Meinard Müller

1. Abstract

Local tempo estimation has never received as much attention in the music information retrieval (MIR) research community as either beat tracking or global tempo estimation. One reason for this may be the lack of a generally accepted definition. We discuss how to model and measure local tempo in a musically meaningful way using a cross-version dataset of Frédéric Chopin's Mazurkas as a use case. In particular, we explore how tempo stability can be measured and taken into account during evaluation. Comparing existing and newly trained systems, we find that CNN-based approaches can accurately measure local tempo even for expressive classical music, if trained on the target genre. Furthermore, we show that different training-test splits have a considerable impact on accuracy for difficult segments.

2. Modeling Local Tempo

Effects of Different Selections

Effects of Different Aggregations





sac = select IBIs, aggregate, and then convert the result to BPM sca = select IBIs, convert to BPM and then aggregate them

3. Tempo Stability

$$c_{var} = \frac{\sigma}{\mu}$$

stability can be defined as the standard Tempo of tempo deviation normalized which is values, equivalent to the coefficient of variation of (sampled) local tempo values, i.e., the ratio between the standard deviation σ and mean μ .

Recordings



4. Experiment

Work

Dataset			
Measures	Beats		
wicasuics	Deats		

Op. 17, 4	132	396	62	
Op. 24, 2	120	360	64	
Op. 30, 2	65	193	34	
Op. 63, 3	77	229	88	
Op. 68, 3	61	181	50	

Overview: Number of measures, beats, recordings for five Chopin Mazurkas [1]. See also http://www.mazurka.org.uk/

Systems

- Böck (BLSTM) [2]
- DeepTemp (CNN) [3]
- DT-Maz_M
- DT-Maz_v

 $DT-Maz_M$ and $DT-Maz_V$ are identical to DeepTemp, but have been trained on different crossvalidation splits [4].



5. Evaluation



6. Conclusions

- Local tempo may be modelled using median aggregated IBIs.
- The local tempo c_{var} may be used as a measure for stability.
- Training on the target genre can lead to convincing results, even for expressive piano music.
- Care must be taken to avoid the "cover song" effect, i.e. overfitting to musical pieces (DT-Maz_V).
- Contrasting models trained on different splits may be used as a tool to identify difficult passages.
- [1] P. Grosche, M. Müller, and C. S. Sapp, "What makes beat tracking difficult? A case study on Chopin Mazurkas," in ISMIR 2010.
- [2] S. Böck, F. Krebs, and G. Widmer, "Joint beat and downbeat tracking with recurrent neural networks," in ISMIR 2016.
- [3] H. Schreiber and M. Müller, "Musical tempo and key estimation using convolutional neural networks with directional filters," in SMC 2019.
- [4] Trained models at https://github.com/hendriks73/tempo-cnn
- [5] H. Schreiber, J. Urbano, M. Müller, "Music tempo estimation: Are we done yet?", in TISMIR 3(1), 2020.



