DMRN+19: DIGITAL MUSIC RESEARCH NETWORK
ONE-DAY WORKSHOP 2024

QUEEN MARY UNIVERSITY OF LONDON
TUE 17 DECEMBER 2024

Can downbeat trackers predict hypermetre?

Jose Alejandro Esquivel de Jesus [1] and Jordan B. L. Smith [2]

[1]Queen Mary University of London, UK, j.esquiveldejesus@qmul.ac.uk
[2]Queen Mary University of London, UK

***Abstract*— Hypermetre is the metric structure between the timescales of downbeats and sections in a piece of music. Hypermetre tracking systems, in principle, face the same difficulties posed by beat and downbeat detection. In this work, we evaluate the success of piggybacking downbeat tracking methods to perform hypermetre tracking. We aim to demonstrate their limitations and the need for further research.**

I. INTRODUCTION

If downbeats are metrically accented beats, then hyperdownbeats are metrically accented downbeats, and each one defines the start of a hypermeasure [1]. Beat and downbeat tracking have been widely researched. Machine learning methods have been used to perform beat and downbeat tracking [2, 3], as well as to predict large-scale structure in songs [4]. Some methods address both timescales in the same model (e.g., [5]), but we are not aware of any that models the in-between timescale of hypermeasures. However, doing it seems worthwhile because it would provide us with improved regularisation of structure analysis that would benefit MIR tasks (e.g., segmentation, AI music generation, and recommendation systems).

II. DATA

We are not aware of any datasets of hypermetre annotations. However, in the McGill Billboard dataset [6] of chord labels, the line breaks in the annotation files appear to indicate the hypermetre. We manually verified this for a set of 26 songs. We introduce the term *Hyper-time-signature* as a way of defining the number of measures to a hypermeasure, similar to the time signature. One common assumption of downbeat trackers is that the time signature remains unaltered in the song [7]. From the annotations in the Billboard dataset, 0.7% of 890 songs exhibit a change in time signature. This value contrasts with 44% of the songs that appear to have a change in hyper-time signature, suggesting a need for further research on hypermetre tracking systems that account for hypermeasure metric changes, given they seem to occur at much higher frequency than measure metric changes.

III. METHOD AND RESULTS

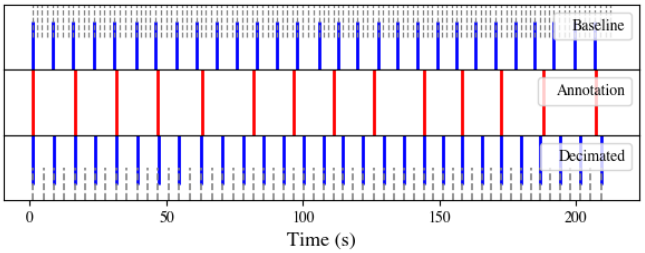

Figure 1: "The Rose" by Bette Middler (McGill Billboard song 0006), as annotated and as predicted by two methods.

We performed an experiment using two methods: the *baseline* and the *decimated*. The *baseline* method serves as a simple baseline for performance floor. We assume 4/4 for the time signature and the hypermeasure-time-signature. This method estimates the beats using spectral flux. Then, downbeats and hyperdownbeats are estimated, assuming they occur every four beats and four downbeats, respectively. The *decimated* method is a first attempt to use an existing beat tracker (madmom) to perform hypermetre tracking. The method uses the madmom downbeattracker to estimate the downbeats locations; then the downbeats are extracted to a new audio file by concatenation and then madmom is re-run on this decimated version to estimate the hyperdownbeats.

In Fig. 1, we can see the results of these methods applied to a pop song. The middle red lines are the human annotations, the blue lines are the hyper-downbeats estimates, and the gray lines are the downbeat estimates for each method. For the *baseline* method, we can observe a misalignment between estimated and annotated times; estimated occur a beat before annotated for the first four red intervals. The fifth red interval is wider than the previous ones because it has a hypermeasure of five measures instead of four. For the *decimated* method, we can see alignment between estimated and annotated hyperdownbeats for the first four red intervals, but this is lost due to the wider fifth red interval.

From these observations, we conclude that hypermetre tracking cannot be reliably performed by only considering beat and downbeat information and that to improve hypermetre tracking, it is necessary to approach hypermetre by accounting for other music characteristics (e.g., harmonic and instrumentation content).

IV. References

[1] H. Krebs, "Hypermeter and Hypermetric Irregularity in the Songs of Josephine Lang." *Engaging Music: Essays in Music Analysis*, 2005.

[2] J. Zhao, G. Xia, and Y. Wang, "Beat Transformer: Demixed Beat and Downbeat Tracking with Dilated Self-Attention," 9 2022. [Online]. Available: http://arxiv.org/abs/2209.07140

[3] T. Kim and J. Nam, "All-in-One Metrical and Functional Structure Analysis with Neighborhood Attentions on Demixed Audio," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, vol. 2023-October. Institute of Electrical and Electronics Engineers Inc., 2023.

[4] O. Nieto, G. J. Mysore, C. I. Wang, J. B. Smith, J. Schluter, T. Grill, and B. McFee, "Audio-Based Music Structure Analysis: Current Trends, Open Challenges, and Applications," pp. 246–263, 2020.

[5] M. Fuentes, B. McFee, H. C. Crayencour, S. Essid, and J. P. Bello, "A Music Structure Informed Downbeat Tracking System Using Skip-chain Conditional Random Fields and Deep Learning," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 5 2019, pp. 481–485. [Online]. Available: https://ieeexplore.ieee.org/document/8682870/

[6] J. A. Burgoyne, J. Wild, and I. Fujinaga, "An expert ground truth set for audio chord recognition and music analysis," in *Proceedings of the 12th International Society for Music Information Retrieval Conference*, Miami, FL, 2012, pp. 633–638.

[7] F. Foscarin, J. Schlüter, and G. Widmer, "Beat this! Accurate beat tracking without DBN postprocessing," 7 2024. [Online]. Available: http://arxiv.org/abs/2407.21658