A COMPARISON BETWEEN TEMPO ESTIMATION ALGORITHMS IN EDM

Rubén Eguinoa Cabrito
Universitat Pompeu Fabra
ruben.eguinoa01@estudiant.upf.edu

ABSTRACT

The tempo estimation of music is a classic task in Music Information Retrieval (MIR). The BPMs (Beat Per Minutes) are the usual way to measure the tempo of the music, which can be described as the speed or the rhythm of the song. On the other hand, the Electronic Dance Music (EDM) is a genre characterized by a strong and consistent beat pattern and it is not usual to have BPM variations during the songs. Many algorithms and models have been developed since the first estimators appearing in the late 90's [8]. In this work, a comparison between four different approaches to global tempo estimation are analysed and assessed for a set of EDM songs.

1. INTRODUCTION

In the definition of any song or musical piece the tempo arises as one of the main characteristic. Likewise, any musician will need to know the velocity of the song to be played, or a DJ will need to know or deduce the BPMs of the next track to make a proper mix. The first automatic tempo estimator appeared in the 90's when Scheirer presented a method to analyse the tempo of musical signals [8]. The article *Music Tempo Estimation: Are We Done Yet?* [11] collects a detailed chronology on the history and developments in the field, a recommended reading.

The electronic music classified as EDM has a strong dependence on the tempo, not only because it is one of its main characteristic but also because traditionally has played an important role in subgenre classification. It is true that nowadays the tempo cliché when producing a specific subgenre track is not that strong, but the dependence on the BPMs is undeniable. This dependence in the BPM makes it necessary a reliable system of tempo estimation.

In this work, the performance of four different global tempo estimators is analysed for the BPM detection in EDM songs, focusing in two different ways to make the analysis: a beat-tracking algorithm [3, 14] and more modern machine learning models [9, 10].

2. METHODOLOGY

For each track included in the dataset the BPM is going to be predicted using two different methods, four models in total. This estimations are going to be compared with the reference tempo provided in the dataset annotations using three different metrics. The process is available in a public GitHub repository [1]

[1] https://github.com/iRubec/tempo-estimators-comparison

2.1 Dataset

Beatport is the leading electronic and EDM online music store [1]. The open source mirdata library offers many datasets and tools involved in MIR [7], among them the Beatport EDM key dataset fulfils the needs of this work, as it provides 1486 two-minute sound excerpts from various EDM subgenres with tempo, single-key labels, comments and confidence levels annotations [5]. When using it for this work (march 2021), only 792 tracks where accessible and 774 had a valid tempo annotation.

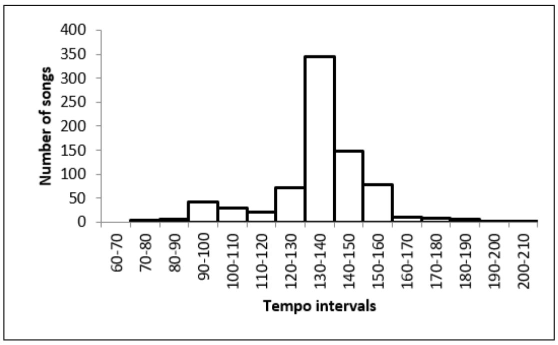

Figure 1. Tempo distribution for the Beatport EDM key dataset with valid annotations.

Size	774
Min BPM	66
Max BPM	195
Avg	126.36
Stdev	17.23

Table 1. Basic stadistics of the Beatport EDM key dataset with valid annotations.

2.2 Tempo estimators

Two different algorithms from the Essentia audio and music analysis, description and synthesis library have been tested [2]: RhythmExtractor2013 [3, 14] and TempoCNN [9, 10].

2.2.1 RhythmExtractor2013

The algorithm extracts the beat positions of the audio input signal and estimates their confidence. Finally it calculates the tempo in BPMs. Among the offered algorithms, the BeatTrackerDegara has been applied [4, 13], which computes a 'complex spectral difference' onset detection function and utilizes the TempoTapDegara beat tracking algorithm to extract beats [4].

2.2.2 TempoCNN

The algorithm uses TempoCNN-based models to estimate the global tempo. Three different models are studied: deepsquare-k16-3, deeptemp-k4-3 and deeptemp-k16-3. They are wrappers that aggregate the predictions generated by TensorflowPredictTempoCNN [9, 10]. The three models work in a similar way: they provide 256 probabilities, one for each BPM value from 30 to 256. The most probable one is the estimation of the tempo of the track. The deeptemp-k4-3 model is expected to be accurate only for rock and pop music, so its performance with EDM will be studied.

2.3 Metrics

To measure how well the models work, it is not enough to measure the exactitude of the data obtained, correct accuracy metrics are also needed. In the 2004 ISMIR conference the accuracy metrics ACC_1 and ACC_2 where established as standards.

ACC_1 computes a 0 or 1 score per track, which means the correctness of the estimation with a 4% tolerance. This tolerance is justified because the Just Noticeable Difference (JND) for music tempo is approximately a 4% [6]. Instead, ACC_2 is more flexible and allows the estimations to be wrong by octave errors (factors of 2, 3, $1/2$ or $1/3$). This measure has created a lot of controversy because it takes as good estimation what should be erroneous.

3. TESTS

To test the four models, the `tempo_eval` evaluation framework has been used [12]. It documents strengths and weaknesses of tempo estimation approaches using a variety of metrics. For each estimation model a document with the results is created by `tempo_eval`.

4. RESULTS

For ACC_1 and ACC_2 similar values appear when a big tolerance is applied. The `k16-3` type models have a way better accuracy up to a 86% in the ACC_1 and a 92% in the ACC_2 for tolerances below 1%. Instead, the other two models do not give acceptable estimations until a 1% of tolerance. See Figures 2 and 3.

Overall results are better for ACC_2, revealing that octave errors are common in the EDM music.

When looking at the exactitude of the estimations, the `deeptemp-k16-3` model gives the better performance, with a 87.47% of exactly predicted BPMs, see Table 2.

Model	ex	%
deepsquare-k16-3	668	86.30
deeptemp-k16-3	677	87.47
deeptemp-k4-3	531	68.60
rhythmExtractor	317	40.96

Table 2. Ratios of exactitud for each model

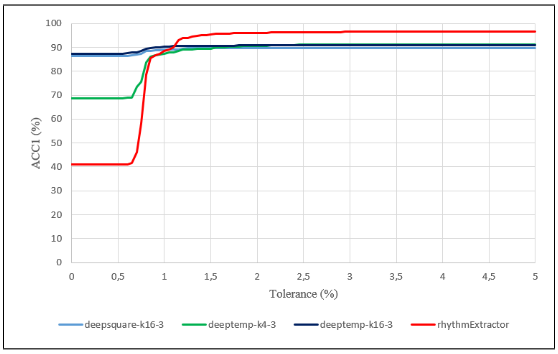

Figure 2. ACC_1 for the four models and different tolerances

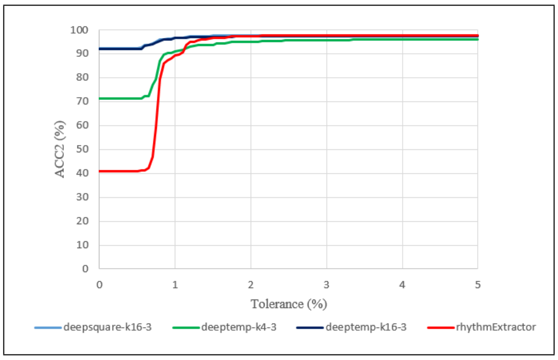

Figure 3. ACC_2 for the four models and different tolerances

5. CONCLUSION AND DISCUSSION

Four different models of global tempo estimation have been analysed and compared. Different metrics on the tempo estimation task have been explained and studied for each model. As EDM is a tempo dependent music genre with a consistent and continuous beat pattern, exact accuracy is asked, so the typical metrics of ACC_1 and ACC_2 with a 4% of tolerance are not tolerable, and the models with a better performance with less than a 1% of tolerance should be taken in account.

None of the studied models have demonstrated a desired 100% of accuracy in the global tempo estimation, but the `deeptemp-k16-3` shows the best overall results.

6. REFERENCES

[1] Beatport, 2021.

[2] Dmitry Bogdanov, Nicolas Wack, Emilia Gómez, Sankalp Gulati, Perfecto Herrera, O Mayor, Gerard Roma,

Justin Salamon, J R Zapata, and Xavier Serra. ESSENTIA: an Audio Analysis Library for Music Information Retrieval. In *International Society for Music Information Retrieval Conference (ISMIR'13)*, pages 493–498, Curitiba, Brazil, 2013.

[3] Norberto Degara Quintela, Enrique Argones Rúa, Antonio Pena, Soledad Torres-Guijarro, Matthew Davies, and Mark Plumbley. Reliability-Informed Beat Tracking of Musical Signals. *IEEE Transactions on Audio, Speech Language Processing*, 20:290–301, 2012.

[4] Norberto Degara Quintela, Enrique Argones Rúa, Antonio Pena, Soledad Torres-Guijarro, Matthew Davies, and Mark Plumbley. Reliability-Informed Beat Tracking of Musical Signals. *IEEE Transactions on Audio, Speech Language Processing*, 20:290–301, 2012.

[5] Ángel Faraldo Pérez. Tonality estimation in electronic dance music: a computational and musically informed examination. *TDX (Tesis Doctorals en Xarxa)*, mar 2018.

[6] Anders Friberg and Johan Sundberg. Time discrimination in a monotonic, isochronous sequence. *Journal of The Acoustical Society of America - J ACOUST SOC AMER*, 98, 1995.

[7] Magdalena Fuentes, Rachel Bittner, Marius Miron, Genís Plaja, Pedro Ramoneda, Vincent Lostanlen, David Rubinstein, Andreas Jansson, Thor Kell, Keunwoo Choi, Tom Xi, Kyungyun Lee, and Xavier Serra. mirdata v.0.3.0, January 2021.

[8] Eric Scheirer. Tempo and beat analysis of acoustic music. *The Journal of the Acoustical Society of America*, 103:588–601, 02 1998.

[9] Hendrik Schreiber and Meinard Müller. A Single-Step Approach to Musical Tempo Estimation Using a Convolutional Neural Network. 2018.

[10] Hendrik Schreiber and Meinard Müller. Musical tempo and key estimation using convolutional neural networks with directional filters. *CoRR*, abs/1903.10839, 2019.

[11] Hendrik Schreiber, Julián Urbano, and Meinard Müller. Music Tempo Estimation: Are We Done Yet? *Transactions of the International Society for Music Information Retrieval*, 3:111, 2020.

[12] Hendrik Schreiber, Julián Urbano, and Meinard Müller. Music tempo estimation: Are we done yet? *Transactions of the International Society for Music Information Retrieval (TISMIR)*, 3(1):111–125, 2020.

[13] J. R. Zapata, M. E. P. Davies, and E. Gómez. Multi-feature beat tracking. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(4):816–825, 2014.

[14] Jose Zapata, Matthew Davies, and Emilia Gómez. Multi-Feature Beat Tracking. *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, 22:816–825, 2014.