

THE EFFECT OF TEMPO TRANSFORMATIONS ON ESSENTIA'S BEAT TRACKING PIPELINES

Background & Motivation



MIR tasks include music recommendation, audio alignment, cover song identification.



Low- (loudness) and high- (genre, mood) level features are extracted from songs.



Same sounding audio clips represented differently when done offline by various people.



Possible reasons : audio codecs, **audio transformations** (audible and inaudible).



Research focuses on temporal (time related) transformations.



Beat is one of the most important building blocks of a musical piece.



Beat detection is used in various multimedia applications: video editing, rhythm games, synchronization.

Research Objectives

Inputs having relevant transformations affect pipelines in a correct way (tempo changes affect beat trackers).

How tempo input transformations affect Essentia's beat extraction algorithms?

What musically meaningful transformations can be applied on inputs?

What data can be used for testing the pipeline?

What measure can be used to test the accuracy of the extractor?

How does the accuracy change with regards to the musical transformations?

What can be deduced from the results and what are their causes?

Methodology & Setup

♫ Musical transformations (experiments):



#1 tempo change of whole song



#3 incremental tempo change throughout song



#2 random tempo changes throughout song



#4 decremental tempo change throughout song

♫ Essentia's *RhythmExtractor2013*; *multifeature* & *degara* methods, window size (40 ms & 70 ms).

♫ Special dataset created from multiple open-source annotated sets (Figure 1).

♫ **F-score**, a measure of test's accuracy (precision and recall).

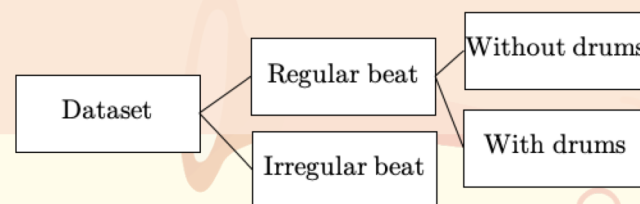


Figure 1. Dataset visualization

Results

- ♫ #1. Tendency where beats are recognized better from the samples with drums in them.
- ♫ #2. 'Regular beat' subsets suffered most (accuracy almost 2 times lower) from random transformation.
- ♫ #3. Speeding up did not drastically affect the scores, small negative impact. Window size plays a role.
- ♫ #4. Slowing down songs had the biggest impact on songs with drums and stable rhythm.
- ♫ **Window size**: crucial parameter, affected the score by up to $\pm 12\%$ with high probability.
- ♫ **Method**: F-scores of multifeature & degara were almost identical, not even 1% difference.

Conclusions & Future Work

- ♫ Experiment results unanimously show, that RhythmExtractor2013 has the highest accuracy on samples which have steady rhythm and drums in them, steady rhythm having more weight.
- ♫ Other parameters – genre and window size – affect the results.

Growing usage of beat trackers in multimedia applications

Growing need for systematic ways to evaluate them



Musical transformations is one way

Generalization not perfect (high SD)

Need a bigger datasets



Spotify Web API

Imperfect audio quality

Using 3rd party libraries for audio processing



Understand their functionality better