

Automated Rhythmic Transformation of Musical Audio

Jason A. Hockman¹, Matthew E.P. Davies², Juan P. Bello³, and Mark D. Plumbley⁴

abstract

Time-scale transformations of audio signals have traditionally relied upon manipulations of tempo. We present a novel technique for the automatic mixing and synchronization between two musical signals. In this transformation, the original signal assumes the tempo, meter, and rhythmic structure of the model signal, while the extracted downbeats and salient intra-measure infrastructure are maintained.

The arbitrary modification of the rhythmic properties of an audio signal offers:

- .. advanced audio editing methods
- .. replication of a MIDI “groove quantization” function
- .. automatic remixing or synchronization between two signals

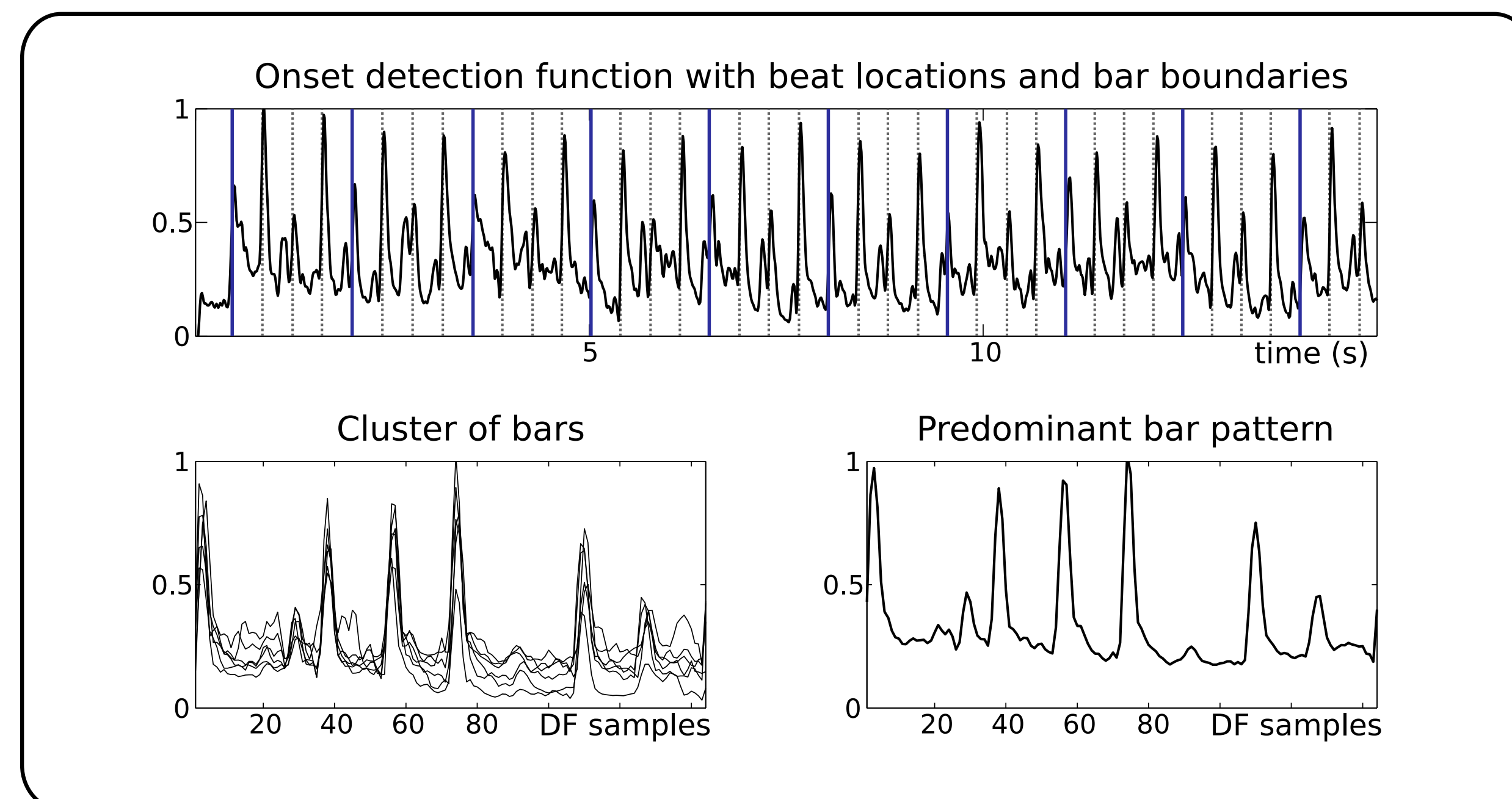
1 rhythmic segmentation

Metrical information

Transform the audio signal into an *onset detection function* from which beat locations and bar boundaries can be found.

Predominant Rhythmic Pattern

Extract the main rhythmic pattern by clustering together each of the bars in the onset detection function signal.



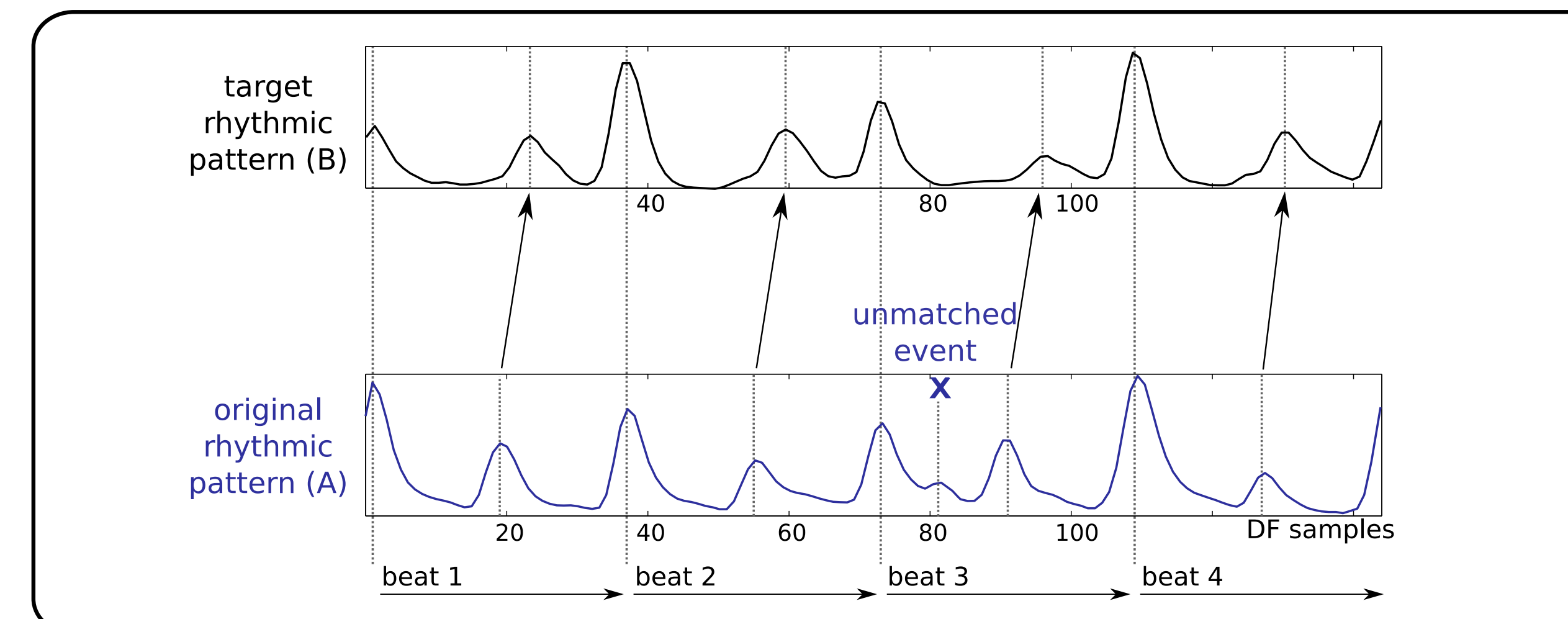
Rhythmic Slices

Peak-pick the predominant rhythmic pattern to give rhythmic pattern events. Rhythmic slices are the sections of the audio signal defined by these pattern events across all extracted bars.

2 rhythmic pattern matching

Signals to match

We select one signal as the *target* (Pattern B) and the other as the *original* signal to be transformed (Pattern A).



To maintain a musically meaningful output, we require that bar boundaries are aligned. For continuity, we must also be able to localize successive beats within the transformed signal.

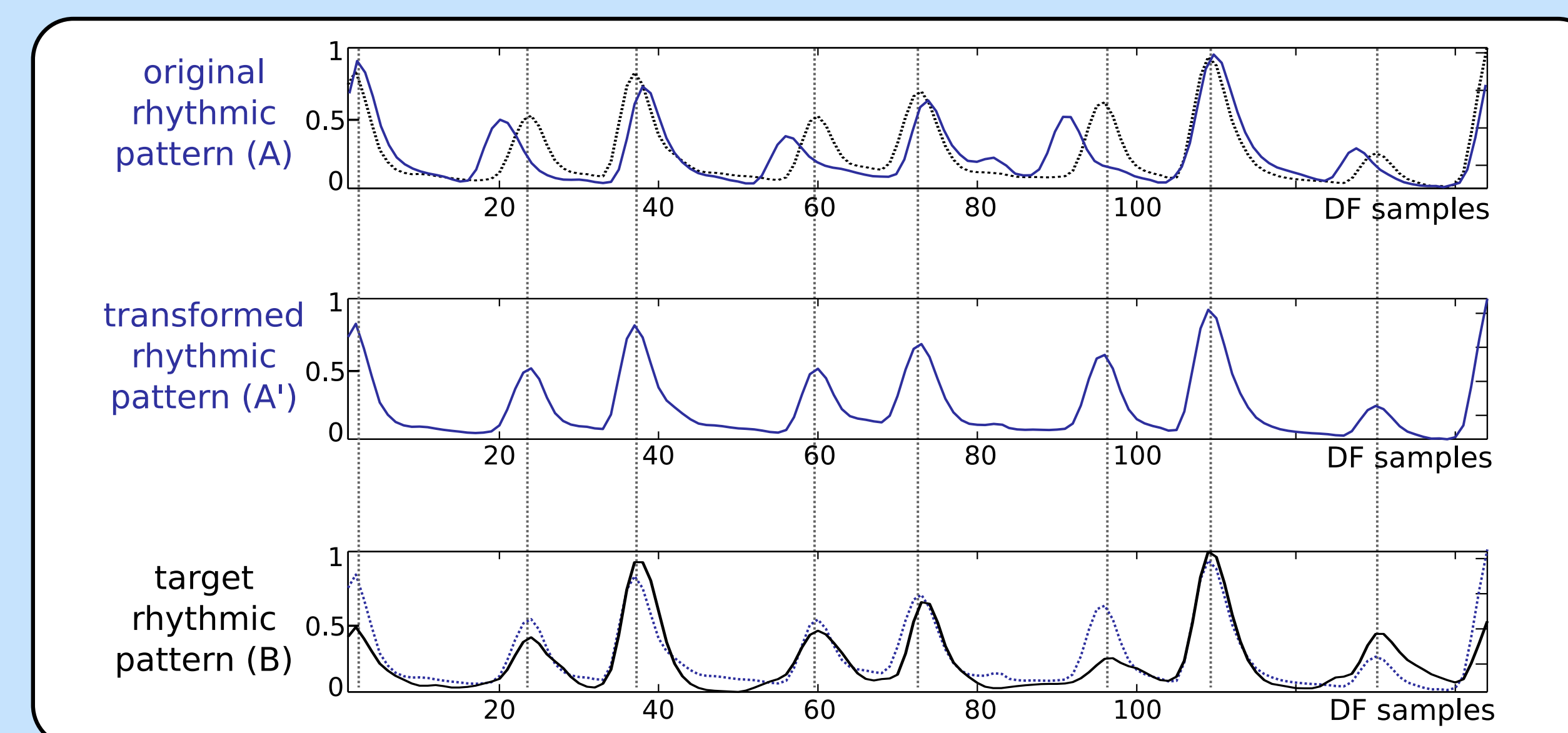
Pattern events per beat

Proceeding *beat-by-beat*, we associate the strongest pattern events with each other. Any excess pattern events in either signal are left unassociated.

3 time-scaling

Rhythmic Slice Processing

Each rhythmic slice in the original signal is now modified by the process of time-scaling to alter its duration to match that of the corresponding event in the target rhythmic pattern. The final output is constructed through the concatenation of each individual rhythmic slice.



4 experimental overview

Database

120 annotated (beats and downbeats) excerpts across 6 genres extracted from an existing database of 523 ballroom dance excerpts. For each genre, a random excerpt is chosen and paired with a random excerpt from another genre.

Classification

We then test the validity of the transformed signal by comparing the results of a genre classifier. Our attempt is to force the misclassification of the original signal as that of the model.

Results

The classification accuracy of the transformed signals (51.5%) demonstrates the manipulative strength of the transform upon classification determination, however is below the classifier performance with the original input genre labels (87.0%).

5 discussion

Challenges

The chief difficulty in the process lies within the initial rhythmic analysis. If this preliminary stage is successful, the resulting transform is surprisingly coherent. A semi-automated system could prevent these errors by allowing the user to adjust the beat and downbeat locations as required.

Future Implementation

We are investigating a real-time version of the transformation for use within production, composition, score-following, and automated accompaniment.



¹jason.hockman@mail.mcgill.ca, ²matthew.davies@elec.qmul.ac.uk, ³jpbello@nyu.edu, ⁴mark.plumbley@elec.qmul.ac.uk