DETECTION OF STRUCTURAL BOUNDARIES IN ELECTRONIC DANCE MUSIC - MIREX 2013

Bruno Rocha
University of Coimbra
bmrocha@dei.uc.pt

Niels Bogaards
Elephantcandy
niels@elephantcandy.com

Aline Honingh
University of Amsterdam
a.k.honingh@uva.nl

EXTENDED ABSTRACT

This document summarizes our submission to the MIREX 2013 Structural Segmentation task and provides pointers to the relevant papers explaining the details of such submission.

We present a method for the detection of structural boundaries in Electronic Dance Music (EDM), of which the main innovations are the addition of a first downbeat detection and the implementation of musically informed rules. The MIR Toolbox [1] is used to perform most of the steps.

Our approach starts by detecting the first downbeat of a track. Many EDM tracks begin with beatless intros and culminate in turning the beat around, a phenomenon that occurs when people perceive a certain metrical structure which is violated later (usually by introducing a beat on the perceived off-beat) [2]. For this reason, the entrance of the bass drum in an EDM track often results in a decisive metrical representation [2]. To detect the first bass drum downbeat, we start by applying a bandpass filter and then compute the global energy of the filtered signal by taking the RMS on non-overlapping windows of 30 seconds, in order to find in which part of the audio file is the beat likely to start (beatless intros usually have low-energy in the low-frequency region). An onset detection is then performed on the thirty seconds window where the energy rises abruptly, leaving us with candidates for the first downbeat. We select the first that exceeds a given threshold and save the previous part as the first segment.

Tempo estimation is then performed in order to detect the duration of a beat. This is important because all features are extracted on beat-related frame lengths. Tempo is computed using the multiplication of two curves: the autocorrelation function computed on the onset detection curve and the spectral decomposition of the onset detection curve. The result is a curve with peaks as indications of the most predominant periodicities found in the track. We then perform peak picking and the highest peak is selected as the tempo of the track. A binary confidence measure telling us about the likeliness that the detected tempo is correct is then derived from the harmonic relation between the found peaks (cf. [3]).

We then calculate the magnitude spectrum for each frame using a Fast Fourier Transform (FFT). The frames are beat-aligned with 87.5% overlap. Subsequently we perform a cepstrum analysis in order to find periodic sequences in the signal. Following Foote [4], we then compute the cosine distance between each possible pair of frames from the cepstrum data to get a self-similarity matrix. Convoluting along the main diagonal of the similarity matrix results in a novelty curve that indicates the temporal locations of significant timbral changes by its peaks. These locations present the segment boundaries that we searched for.

Finally we propose a set of heuristic rules to align the obtained novelty peaks with the beats. In line with Butler [2], who explains how EDM structure relies on sequences of 8 or 16 bars of 4 beats, these heuristics determine an asymmetric weight to these bars.

Although this method was created specifically for EDM, results obtained during this research show that the algorithm performs well when applied to pop music [3], suggesting that the structural boundaries underlying EDM follow the same principles as the ones in pop music.

For the final labeling step, we extract the features explained in [3] and cluster the segments according to their similarity, as measured by the combination of these features.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.
© 2012 International Society for Music Information Retrieval.

ACKNOWLEDGMENTS

This research has been funded by the Center for Digital Humanities (http://cdh.uva.nl/) and the Netherlands foundation of Scientific Research (NWO), grant no. 639.021.126.

REFERENCES

[1] O. Lartillot and P. Toiviainen: “A Matlab toolbox for musical feature extraction from audio,” *Proceedings of the DAFx*, 2007.

[2] M. Butler: *Unlocking the Groove: Rhythm, Meter, and Musical Design in Electronic Dance Music*, Indiana University Press, 2006.

[3] B. Rocha, N. Bogaards, & A. Honingh.: "Segmentation and Timbre Similarity in Electronic Dance Music," *Proceedings of the SMC*, 2013.

[4] J. Foote: “Automatic audio segmentation using a measure of audio novelty,” *Proceedings of the ICME*, 2000.