UNSUPERVISED DETECTION OF STRUCTURAL CHANGES IN ELECTRONIC DANCE MUSIC

Bruno Rocha[1], **Niels Bogaards**[2], **Aline Honingh**[1]

[1]Institute for Logic, Language and Computation, University of Amsterdam, The Netherlands

[2]Elephantcandy, Amsterdam, The Netherlands

{B.M.MachadoRocha, A.K.Honingh}@uva.nl, niels@elephantcandy.com

EXTENDED ABSTRACT

Music structure segmentation is a relevant topic in music information retrieval as it is both useful to the analysis of musical structure and a means to improve performance in other tasks [1], such as audio editing in recording work-flows [2].

Recent years have seen Electronic Dance Music (EDM) gaining access to mainstream charts. However, it is still a neglected genre in music research with just a few articles published in journals and conferences. One exception is Butler [3], who claims that "EDM presents a distinctive overall formal structure in addition to its characteristic instrumentation".

We present a method for the location of structural changes in EDM. Our approach starts by detecting the first downbeat of a track. Tempo is then computed using a standard tempo detection algorithm [7].

We then calculate the magnitude spectrum for each frame using a Fast Fourier Transform (FFT). The frames are beat-aligned with 87.5% overlap. Contrarily to most common approaches, we perform a cepstrum analysis in order to find periodic sequences in the signal.

Based on the approach by Foote [5], we then compute the cosine distance between each possible pair of frames in the cepstrum data to get a self-similarity matrix. Convoluting along the main diagonal of the similarity matrix using a Gaussian checkerboard kernel yields a unidimensional linearly normalized novelty curve that indicates the temporal locations of significant textural changes. We use a kernel size of approximately 30s. Justifications for this value can be found in [6]. Positions of prominent peaks of the novelty curve are selected as candidates for segment boundaries.

Finally we propose a set of heuristic rules to align the obtained novelty peaks with the beats (see Figure 1). In line with Butler [4], who explains how EDM structure relies on sequences of 8 or 16 bars of 4 beats, these heuristics apply an asymmetric weight to these bars.

Table 1 shows the results obtained for different datasets using the same parameter settings. It performs well on an in-house EDM dataset[i]. Although this method was created specifically for EDM, results on the RWC Pop dataset [8] are in line with the best performing algorithms submitted to MIREX 2012[ii]. This means that structural changes in pop music might have the same periodicity as in EDM. However, this method does not reach high performance on the Eurovision dataset. An explanation for this might be that pop music is mixed with traditional music from European countries in this song contest.

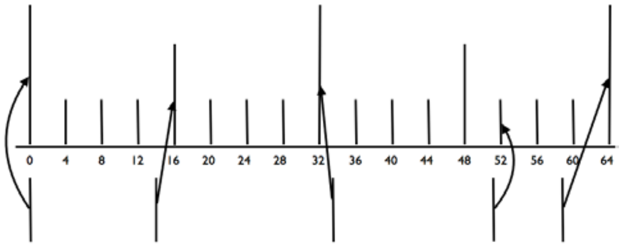

Figure 1: Timeline is shown in beats (0 is the first beat detected). Heuristic rules dictate an asymmetric weight towards the 8th and 16th bars.

Dataset	P0.5s	R0.5s	F0.5s	P3s	R3s	F3s
EDM	45.14	56.20	50.07	66.12	83.81	73.92
RWO [8]	31.76	28.94	30.29	71.07	65.76	68.31
RWQ [9]	33.50	31.29	32.35	68.17	63.82	65.92
EUR [9]	13.48	13.78	13.63	44.57	46.59	45.56

Table 1: Boundary retrieval precision rate (P), recall rate (R) and F-score (F) with two tolerance windows: ±0.5 seconds and ±3 seconds. Three annotated datasets were used: in-house (EDM), RWC (original (RWO) and Quaero (RWQ) annotations) and Eurovision (EUR).

To conclude, we have presented a system that detects structural changes based on timbre. The success of the method is in agreement with Butler's suggestion that EDM rejects harmony as primary musical parameter [3].

REFERENCES

[1] M. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney: "Content-based music information retrieval: current directions and future challenges," *Proc. of the IEEE*, 96(4): 668–696, 2008.

[2] G. Fazekas and M. Sandler, "Intelligent editing of studio recordings with the help of automatic music structure extraction," *Proc. of the AES Convention*, 2007.

[3] M. Butler: "Turning the beat around: reinterpretation, metrical dissonance, and asymmetry in electronic dance music," *Online Journal for the Society of Music Theory*, 7(6), 2001.

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

[5] J. Foote: "Automatic audio segmentation using a measure of audio novelty," *Proc. of the ICME*, 2000.

[6] J. Serrà, M. Müller, P. Grosche, and J. Arcos: "Unsupervised detection of music boundaries by time series structure features," *Proc. of the AAAI*, 2012.

[7] O. Lartillot and P. Toiviainen: "A Matlab toolbox for musical feature extraction from audio," *Proc. of the DAFx*, 2007.

[8] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka. "RWC Music Database: Popular, Classical, and Jazz Music Databases," *Proc. of ISMIR*, 2002

[9] F. Bimbot, O. Le Blouch, G. Sargent, and E. Vincent, "Decomposition into autonomous and comparable blocks: a structural description of music pieces," *Proc. of ISMIR*, 2010.

---

[i] The list of songs can be found at http://brunoaudio.weebly.com/research.html

[ii] Algorithm performance summaries can be consulted at http://nema.lis.illinois.edu/nema_out/mirex2012/results/struct/mrx10_2/summary.html and http://nema.lis.illinois.edu/nema_out/mirex2012/results/struct/mrx10_1/summary.html