UNSUPERVISED DETECTION OF STRUCTURAL CHANGES IN ELECTRONIC DANCE MUSIC

Bruno Rocha¹, Niels Bogaards², Aline Honingh¹

¹Institute for Logic, Language and Computation, University of Amsterdam, The Netherlands
²Elephantcandy, 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 workflows [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 towards the 8th and 16th bars.

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


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 Quaoer (RWQ) annotations) and Eurovision (EUR).

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

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.

¹The list of songs can be found at http://brunoaudio.weebly.com/research.html
²Algorithm performance summaries can be consulted at http://nema.lis.illinois.edu/nema_out/mirex2012/results/struct/mrx10_1/summary.html and http://nema.lis.illinois.edu/nema_out/mirex2012/results/struct/mrx10_2/summary.html