TEXTURAL LAYERS AND POLYPHONIC TIMBRE LINKS IN ELECTRONIC DANCE MUSIC

María Marchiano
Laboratorio para el Estudio de la Experiencia Musical (FBA-UNLP)
marchiano.maria@gmail.com

Isabel Cecilia Martínez
Laboratorio para el Estudio de la Experiencia Musical (FBA-UNLP)
isabelmartinez@fba.unlp.edu.ar

Javier Dameśon
Laboratorio para el Estudio de la Experiencia Musical (FBA-UNLP)
javier.dameson@gmail.com

ABSTRACT

Polyphonic timbre refers to the overall timbral mixture in a music signal. Some of their most salient acoustic dimensions is the Sub-Band Flux. Our hypothesis is that there is a temporal alignment between the Sub-Band Flux changes and the perceived textural layers' onsets and offsets in electronic dance music (EDM). The study had two methods: (i) an experiment in which 15 professional musicians were asked to record in Reaper every perceived textural layer's onset or offset of 11 EDM's tracks, (ii) and the estimation of the transitional temporal data between homogeneous and successive states of Sub-Band Flux from sound signals. Both data sets were correlated. We found some temporal coincidences between the Sub-Band Flux changes and the perceived textural layers' onsets and offsets. Lowest spectral flux's bands show significant correlations with perceived texture, and it could mean that timbre and texture share a mutual acoustical background.

1. BACKGROUND

We perceive timbre as a multidimensional representation [1]. Nowadays, the study of timbre tries to find the acoustical features that afford our multidimensional perception.

There are two main categories of timbre based on different ideas of timbre blends: individual or monophonic timbres, and emergent or polyphonic timbres [2]. Monophonic timbre has been the main subject of study for a long time, and some of their acoustical features with perceptual correlations are spectral centroid, spectral flux, log-attack time, roughness and the Mel-Frequency Cepstral Coefficients (MFCC), among others. However, monophonic timbres are almost non-existent in our daily auditory experience. On the contrary, our sonic and musical environment is formed by polyphonic timbre.

Polyphonic timbre is defined as the overall timbral mixture in a music signal [3]. Its acoustical dimensions are currently being explored. Alluri [2] has analyzed a great quantity of acoustical features on a set of musical stimuli, to establish their relevance in the auditory experience of polyphonic timbres. One of them is the Sub-Band Flux feature. This feature "represents the fluctuation of frequency content in ten octave-scaled bands of the spectrum. (…) For each of the ten channels the spectral flux was calculated as the Euclidean distance between successive amplitude spectra" [2]. Alluri [4] found that some bands show strong correlations with timbral semantic descriptions of Indian music. The three lower bands (0-200 Hz) are related with the sensation of "fullness". Hartmann, Lartillot and Toivianen [5, 6] also find that the Sub-Band Flux plays a role in the perceptual segmentation of music. Unlike Alluri, they do not work with each band separately, but with a blend of all spectral flux's bands.

We think that the Sub-Band Flux might be a salient acoustical feature of polyphonic timbre in other music styles and in different types of perceptual tasks.

Beyond that the concept of polyphonic timbre does not refer to the textural aspect of music, but relates to the emergence of polyphonic timbres in non-monophonic textures. According to Fessel [7], the textural types are defined by five features: layers, homogeneity, linearity, accent coincidence, and attack divergence.

In electronic dance music (EDM), the layers are the most relevant feature for the definition of textural types. Normally, there is not a lead voice in the EDM, because all layers have relatively the same importance [8]. This characteristic produces some specific textural types, as mass texture, extended heterophony, paraphony, and rhythmic or timbral polyphony, and the EDM genres are partly differentiated by their textural features [9]. According to Anzil, audio-tactile sensation is a key feature in composition and in multisensorial perception of timbre in the EDM parties. This sensation is clearer in the low frequency zone.

2. HYPOTHESIS

People are perceptually sensitive to textural layers' onsets and offsets in EDM. These perceived transitions are temporarily aligned with Sub-Band Flux changes. We anticipate that the most perceptually relevant textural layers' onsets and offsets will be temporarily aligned with the most pronounced acoustic changes.

3. AIMS

To find temporal correlations between timbral and textural changes, with the long-term perspective of developing a model about these linkages.

4. METHOD

4.1 Experiment

4.1.1 Subjects

15 professional musicians participated in the experiment (MA=31; SD=6.9; average 15.8 years of music studies in institutions). All subjects had heard of EDM, but only 4 had been or currently were users of EDM parties.

4.1.2 Stimuli

11 EDM's tracks were used on the experiment (2:45 minutes average duration per track) [10].

The selection was based in the variety of EDM's genres, with different textural and timbral qualities. The genres of selected tracks were deep and electro house, techno, trap, acid and trance. This selection was made by auditory analysis of the authors.

4.1.3 Procedure

The experiment was done on Reaper v.5 with two Mackie mk3 mr6 monitors. We provided 11 Reaper projects randomly ordered (one for each track). Participants were asked to record every perceived textural layer's onset and offset of EDM's tracks by pressing the letter "m" of the computer keyboard. In order to get temporarily accurate marks, this was not a real-time task: participants could listen to each track as often as they needed, and delete, add or change the temporary location of marks.

The task involved an analytical listening, where participants had to determine what was a textural layer in the music they were working on. We asked them to make a decision using a single criteria for each track, giving the example: "if you make a mark for a little sound that you would consider a textural layer, you must mark all layers' onsets and offsets of that level of significance. This level of significance has to emerge from the music and could change from track to track".

After the task, we asked participants if the textural layer's onsets and offsets were perceived as musically relevant.

4.2 Computational modelling

All analysis were made on Matlab R2015a. We use MirToolbox1.7 for the feature extraction and novelty detection [11].

4.2.1 Sub-Band Flux extraction

Sub-Band Flux was computed by firstly filtering the sound signal in 10 spectral bands, and secondly extracting the spectral flux for each band. The analysis window normally used for Sub-Band Flux extraction is 0.025 seconds with 50% overlapping. Given the subsequent analysis, we used a 1 second window without overlapping (see *Novelty Detection*).

We used two sets of Sub-Band Flux data. On the one hand, a single temporal series of Sub-Band Flux was computed by combining the data of all bands. This series bears the Sub-Band Flux information of the complete spectrum. We will call it *general Sub-Band Flux data*, to differentiate it from the other set. On the other hand, spectral flux data of each band was used separately.

4.2.2 Novelty Detection

The novelty function detected the dissimilarity between homogeneous and successive states, and the relative importance of these temporal transitions. The peaks of the resulting novelty curve represent the moments of structural change in the sound. These segmentation points are often correlated with perceptual tasks results [12], to evaluate the perceptual relevance of the acoustical structure of a musical signal.

There are several approaches to novelty detection [13, 14, 6]. We decided to use the multi-granular method proposed by Lartillot et al. [14], because its algorithm can detect homogeneous states of different temporal lengths. The computation of the novelty curve in this approach is attached to the previous feature window analysis. We processed the signal with different window sizes to determine which have a more adjusted temporal alignment with the perceived layer's onsets and offsets. In order to get an acoustical data to correlate with perceptual task results and considering their possible temporal imprecisions, we decided that a 1 second window without overlapping was the better choice.

5. RESULTS

5.1 Perceptual Task

All participants of the experiment showed perceptual sensitivity to textural layers' onsets and offsets, and 93.3% judged that this textural layer behavior was relevant in the music of the task.

Frequency of perceived layers' onsets and offsets were calculated by counting the number of all participants' marks made in the same second (frequency range: 1 to 15). The maximum frequency represents the 3.83% of total marks of the experiment (SD of all tracks frequency=0.28, normalized).

5.2 Perceptual and Acoustic Correlations

Initially, we searched for temporal coincidences between the higher values or peaks of the novelty curve of general Sub-Band Flux and all marks of the perceptual task (of all participants and all tracks). There was a total of 47.12% coincidence, among all the tracks. Percentage of coincidences for every track is shown on Table 1.

Track	Temporal coincidence (%)	Correlation Coefficients
1	44.4	0.58
2	24	0.5
3	51.6	0.43
4	43.6	0.07
5	59.4	0.36
6	34.8	0.42
7	40.9	0.5
8	39.4	0.39
9	52.6	0.5
10	46.5	0.54
11	47.6	0.38

Table 1. Percentage of temporal coincidences between perceptual marks of experiment and novelty peaks of general Sub-Band Flux; correlation coefficients between marks' frequency and novelty values

We ran two set of correlations. In both, we searched for the novelty values of Sub-Band Flux for the temporal locations of all marks made by every participant in the perceptual task, and correlated the mark's frequency and their corresponding novelty values.

In the first set of correlations, we used the general Sub-Band Flux data, which includes information of complete spectrum. Results show significant correlations for some tracks and low correlation coefficients for others (medium-high significance = 45.5%; medium-low significance = 45.5%; no correlation = 9%) (Table 1).

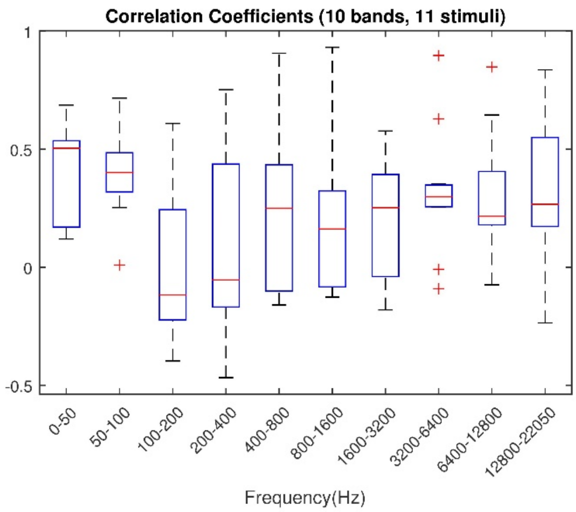

Figure 1. Box plot of correlation coefficients of each spectral flux's band between their values and the frequency of perceptual marks, in the 11 stimuli.

In the second set, we correlated the marks' frequency of the perceptual task and the values of each separately spectral flux band for each perceptual mark's temporal location. Correlation between frequency of all marks and the temporal series for the separately spectral flux's bands did not show significant results. We thought that it is possible that the perception of textural layer's onsets and offsets is linked to acoustic changes in some part of spectrum, as the Sub-Band Flux feature itself proposes. Thus, the perceptual marks not related to a high value of a spectral flux's band, could be related to another.

To confirm this, we started searching the higher peaks of the novelty curves of all spectral flux's bands. Then, we selected the perceptual marks that were temporarily coincident with any of the novelty's higher peaks, and got 10 different sets of marks (one per spectral flux's band). Finally, we correlated the selected marks' frequency and their novelty values. This method increases the correlation significance: spectral flux's bands 1 and 2 show significant correlation values (band 1's mean= 0.5; band 2's mean= 0.4) (Figure 1).

6. DISCUSSION

The present study found some links between the perceived textural layer's onsets and offsets, and Sub-Band Flux changes in EDM. However, these links are not as strong as we expected.

The analysis of a general Sub-Band Flux's novelty –i. e., novelty of the blending of all band's data in a single curve- shows a clear relationship with the perception of textural layer's changes. This connection is not so strong in a more detailed analysis. Low correlation coefficients for each spectral flux's band make us think that its changes are not so relevant to the layer's onsets and offsets perception. The links between changes in Sub-Band Flux and texture could be occurring at an acoustic and perceptual global level.

On the one hand, we think it is likely that other acoustic features are involved in the perception of texture layers. This could be explained through a joint analysis of a bigger set of features. On the other hand, it is also possible that this unstable link between Sub-Band Flux and texture is an EDM's stylistic characteristic. It would be necessary to include a more stylistically varied set of stimuli in a future study.

However, there are two spectral flux's bands which show significant correlations with the perceived layers: band 1 (0-50 Hz) and band 2 (50-100 Hz). Alluri had found that these lowest bands are related to the "fullness" descriptor of timbre sensation. This result not only contributes to the idea of the perceptual relevance of the lowest Sub-Band Flux behavior, but also allows to think of a possible shared acoustical background for timbre and texture perception.

Moreover, the lower frequencies have special treatment in the EDM and they are audio-tactilely perceived in a club with a specific kind of amplification. It is interesting to notice that even in a completely different acoustic environment, the low frequencies seem to be more perceptually relevant regarding texture than the rest of the spectrum.

7. REFERENCES

[1] McAdams, S. (1999). Perspectives on the contribution of the timbre to musical structure. *Computer Music Journal, 23*(3), 85-102.

[2] Alluri, V. (2012). *Acoustic, Neural, and Perceptual Correlates of Polyphonic Timbre* (Doctoral Thesis). University of Jyväskylä: Jyväskylä.

[3] Aucouturier, J., Pachet, F., and Sandler, M. (2005). "The Way It Sounds": Timbre Models for Analysis and Retrieval of Music Signals. *IEEE Transactions on multimedia, 7*(6), 1028-1035.

[4] Alluri, V. and Toivianen, P. (2009). Exploring Perceptual and Acoustical Correlates of Polyphonic Timbre. *Music Perception, 27*(3), 223–241

[5] Hartmann, M., Lartillot, O. and Toivianen, P. (2017a). Interaction features for prediction of perceptual segmentation: Effects of musicianship and experimental task. *Journal of New Music Research, 46*(2), 156-164.

[6] Hartmann, M., Lartillot, O. and Toivianen, P. (2017b). Musical Feature and Novelty Curve Characterizations as Predictors of Segmentation Accuracy. *SMC 2017: Proceedings of the 14th Sound and Music Computing Conference 2017,* 365-372.

[7] Fessel, P. (1997). Hacia una caracterización formal del concepto de textura. *Revista del Instituto Superior de Música, 5*, 75-94.

[8] Butler, M. (2003). *Unlocking the groove: Rhythm, meter and musical design in electronic dance music.* Indiana University: Indiana.

[9] Anzil, I. (2016). *Sensación táctil y audiotáctil en la música. El caso de las músicas electrónicas utilizadas para el baile social en locales de baile de la Ciudad Autónoma de Buenos Aires y alrededores* (Doctoral Thesis). Universidad Nacional de La Plata, La Plata.

[10] Varios. [María March]. (2017, 12, 22). EDM SysMus [Archivo de video]. Recuperado de https://www.youtube.com/playlist?list=PLWWthKwdwpnVrpumIDccjJw-utN_EYhn2

[11] Lartillot, O. and Toivianen, P. (2007). A Matlb Toolbox for Musical Feature Extraction from Audio. *Proceedings of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*. Bordeaux, France.

[12] Hartmann, M. (2017). *Modelling and Prediction of Perceptual Segmentation* (Doctoral Thesis). University of Jyväskylä: Jyväskylä.

[13] Foote, J. T. (2000). Automatic Audio Segmentation Using a Measure of Audio Novelty. *Proceedings of IEEE-ICME, 1*, 452–455.

[14] Lartillot, O., Cereghetti, D., Eliard, K. amd Grandjean, D. (2013). A Simple, high-yield Method for Assessing Structural Novelty. *Proceedings of the 3rd International Conference on Music & Emotion (ICME3).* Jyväskylä, Finland.