21ST CENTURY ELECTRONICA: MIR TECHNIQUES FOR CLASSIFICATION AND PERFORMANCE

Dimitri Diakopoulos, Owen Vallis, Jordan Hochenbaum, Jim Murphy, Ajay Kapur

California Institute of the Arts
Valencia, CA USA
{ddiakopoulos,jamesmurphy}
@alum.calarts.edu

New Zealand School of Music
Wellington, NZ
{ajay.kapur,vallisowen,hochenjord}
@nzsm.ac.nz

ABSTRACT

The performance of electronica by Disc Jockys (DJs) presents a unique opportunity to develop interactions between performer and music. Through recent research in the MIR field, new tools for expanding DJ performance are emerging. The use of spectral, loudness, and temporal descriptors for the classification of electronica is explored. Our research also introduces the use of a multi-touch interface to drive a performance-oriented DJ application utilizing the feature set. Furthermore, we present that a multi-touch surface provides an extensible and collaborative interface for browsing and manipulating MIR-related data in real time.

Keywords: Electronica, Electronic Dance Music, Genre Classification, User Interfaces, DJ, Multi-touch.

1. INTRODUCTION

Electronic dance music, often referred to as Electronica, is an overarching collection of genres that focus predominately on rhythmic motifs & repeating loops. A task of the electronica DJ is to compile a set-list of music for performance. Additionally, DJs are always looking for ways to expand the interactivity of their performances through the use of new tools. The primary goal of this work is to give the modern, digital DJ access to a wider range of performance options using MIR techniques such as feature extraction, genre classification, and clustering. Combined with advances in tabletop computing, these techniques have made it possible to add a layer of interactivity to automatic playlist generation.

In the following section we detail related work on music features and electronic performance interfaces including recent work in tabletop computing. In the remainder of the paper we discuss our feature extractors, genre classification results and the interface that we developed to enable DJs to interact with those results to create set-lists. We conclude the paper with a discussion of future work in interactive MIR powered DJ applications and tabletop computing.

2. RELATED WORK

Our work draws on a wide array of related research ranging from musical descriptors and novel performance interfaces to recent applications in tabletop computing. By synthesizing these related but disparate areas of research, we enable new performance experiences for individual and group DJs to create and modify set-lists in real time.

Genre classification can be accomplished using a range of signal features and algorithms. For electronica in particular, features and patterns such as rhythm, tempo, periodicity, and even use of panning have been explored in the literature [1-3].

For DJs specifically, the use of interfaces to retrieve musically relevant material in performance has included query-by-beat-boxing [4], and query-by-humming [5]. Retrieval using both traditional and non-traditional instruments and interfaces has been explored by [6]. Other research in the academic arena for enabling DJ performance includes AudioPad [7], and Mixxx [8]. Although we take influence in these interfaces for retrieval, our work wishes to explore a browsing paradigm using similar creative interfaces.

In the commercial sector, Stanton's Final Scratch[1] enables DJs to use a physical controller to manipulate and mix digital music, while Native Instruments' Traktor[2] is a software-only solution for DJ performance. Ableton's flagship software, Live[3], has been increasingly used to enable DJs to use their own pre-composed music in live performance through the synchronized playback of different audio loops, known as clips.

A multitude of literature on tabletop computing & interfaces exists. The Reactable team was one of the first groups to directly apply both tangible and multi-touch interaction to the performance of music [9], followed by others including the earlier referenced *AudioPad,* which is also a tangible interface. More recently*, MarGrid,* a UI for the browsing of a digital music collection using Self-Organizing Maps has been examined using a tabletop interface [10]. The use of Self Organizing Maps (SOM) for visualizing feature data has also been previously covered by [11], [12]. In addition, although not performance-oriented, *MusicSim* presents an interesting combination of audio analysis and music browsing in an interactive computer-based interface [13].

Our aim here is to expand on these efforts by introducing the use of a multi-touch surface in a way that is both intuitive and collaborative. The use of Self Organizing

[1] http://www.stantondj.com/
[2] http://www.native-instruments.com/
[3] http://www.ableton.com/

Maps represents a useful way of organizing features for visualization, on-top of which many real-time interactive applications are possible.

3. DATA COLLECTION

For our experiments, six genres across the spectrum of electronic music were selected for their diverse characteristics and wide-spread popularity.

One hundred 2 to 8 minute prototypical tracks were sliced at random into single 30-second chunks for each genre. Our dataset contains at least 20 distinct artists in each genre; tracks were not chosen on the perceived genre of the composing artist, but a human baseline analysis by the authors. In total, there are 600 30-second clips, each in a stereo 44.1 kHz PCM-encoded file format. All files were normalized before experimentation.

3.1 Genre Definitions

Many subgenres fall beneath the umbrella term of electronica—this paper examines six of the most broad & popular genres commonly played by DJs: intelligent dance music (IDM), house, techno drum and bass (DnB), trance, and downtempo. A brief description of them is as follows:

IDM distinguishes itself by its heavy use of complex meter, sophisticated and often sporadic percussive elements, and varying use of syncopation. IDM carries with it a rich harmonic and melodic palate borrowed from many genres. Tempos typically range from 150-180 BPM. Notable artists in the genre are Aphex Twin, Squarepusher, and Autechre. IDM may sometimes be referred to as Glitch music.

House music makes use of the common 'four-on-the-floor' rhythm pattern consisting of a steady kick drum on each downbeat in a 4/4 meter. Defining characteristics involve offbeat open hi-hat patterns and snare or claps on the two and four of every bar. Harmonic content and instrumentation is often borrowed from Disco genres. Tempos usually range from 115 to 135 BPM. Daft Punk, Thomas Bangalter, and Alan Braxe are popular artists in the genre.

Techno uses minimal melodic ornamentation, relying more on bass riffs and polyrhythmic drums layered over a common four-on-the-flour kick drum. The rhythmic elements in techno are often the defining features of the song, with percussive grooves and riffs taking precedence over more traditional melodic and harmonic structure. Significant artists include Derrick May, Richie Hawtin, and Robert Hood.

DnB makes heavy use of "break beat chopping,"—the re-sequencing of drum hits from other previously recorded material. DnB is often composed above 160 BPM, with characteristic bass lines moving at half the tempo. Goldie and Pendulum are both well-known artists.

Trance distinguishes itself by employing thick, complex harmonic components, leaving little room for the complex rhythmic structures found in other similar genres. Trance often makes use of arpeggios, drum rolls, and long crescendos of synthesizers. The genre is composed around 140 BPM. DJ Tiesto, Ferry Corsten, and Sasha are popular artists within the Trance genre.

Downtempo employs lush harmonic textures and groove-oriented percussion. Tempos are characteristically low, ranging from 60 to 90 BPM. Boards of Canada, Air, and Bonobo are well-known artists within the genre.

4. AUDIO ANALYSIS AND CLASSIFICATION

Audio analysis was performed using the ChucK audio programming language [14]. Our results are based on a two-second (88200 sample) Hann window, resulting in 15 8-dimensional vectors for each audio clip. In addition to being written to disk for further analysis, the raw data was also sent over networked protocol (OSC[1]) into Processing[2], a visuals-oriented programming language. The process of visualizing the data using Processing is later presented in Section 5. Before application development could begin, a central concern was to uncover a feature-set that could accurately classify electronica. We follow with a description of the eight features used in our experiments.

4.1 Spectral Features

- Centroid, the centre of mass of the spectrum;
- Flux, the change in spectral energy across successive frames;
- Rolloff, the frequency below which resides 85 percent of a spectrum's energy.

4.2 Loudness Features

- RMS, the amplitude of a window;
- Panning, a coefficient used to describe the weight of the signal in either the left or right channels [3];
- Panning Delta, change in the panning coefficient across successive windows [3].

4.3 Temporal Features

- Number of Bass Onsets, an integer representing the number of peaks ('Beats') detected in a window;
- Average Inter-onset Time, a basic feature to describe the periodicity of the beats across a window.

4.4 Classification

Four separate classifiers were run on all six classes, and also on a smaller set of four classes. All experiments were performed utilizing a 10-fold cross-validation method in the Weka machine learning environment [15].

A k-Nearest Neighbour classifier (IBk) gave the best overall result, resting at a 75.2% classification rate across the six classes (16.7% baseline accuracy). Table 1 shows the confusion matrix for this experiment.

---

[1] http://opensoundcontrol.org/
[2] http://www.processing.org/

As Table 1 illustrates, the k-NN classifier had trouble distinguishing between IDM & DnB, and House & Techno. This is most likely attributed the sporadic percussive elements found in IDM & DnB, and very similar tempos found in House & Techno. Another experiment was run omitting IDM and House, resulting in a superior 87.0% classification rate. In the context of real-time performance and playlist generation, the omission was the result of IDM and House being considerable similar to genres already being classified. In favor of omitting any single pair of the confused genres, one of each was left out. The confusion matrix of this experiment is shown in Table 2.

Other classifiers used in testing were a C4 Decision-Tree (J48), a backpropagation Artificial Neural Network (MultiLayerPerceptron), and a Support Vector Machine (SMO). More details on these classifiers can be found in [15, 16]. Table 3 lists the accuracy of the four different classifiers using both the six-class and four-class datasets.

The exclusion of the two panning features and average inter-onset for the 6 and 4-class datasets using k-NN reduced classification accuracy by 7.70% and 4.85% respectively, indicating that both temporal and panning features moderately improved classification. Given the distinct tempos and production values between electronica genres, higher-level features using both tempo and panning should be considered an important facet of future classification experiments.

	Idm	**Tno**	**Dnb**	**Hse**	**Trn**	**Dtm**
Idm	**0.60**	0.02	0.16	0.07	0.02	0.12
Tno	0.02	**0.84**	0.03	0.07	0.04	0.01
Dnb	0.10	0.03	**0.72**	0.05	0.05	0.04
Hse	0.04	0.05	0.06	**0.78**	0.04	0.03
Trn	0.01	0.04	0.05	0.05	**0.82**	0.02
Dtm	0.13	0.01	0.06	0.05	0.02	**0.74**

Table 1 Confusion matrix, in percent, for the 6-class k-NN classifier

	Techno	**Dnb**	**Trance**	**Dtempo**
Techno	**0.90**	0.05	0.05	0.01
Dnb	0.04	**0.84**	0.06	0.06
Trance	0.05	0.06	**0.87**	0.02
Dtempo	0.01	0.10	0.02	**0.87**

Table 2 Confusion matrix, in percent, for the 4-class k-NN classifier

	IBk	**J48**	**MLPercept.**	**SMO**
6 Class	0.75	0.66	0.60	0.58
4 Class	0.87	0.82	0.81	0.79

Table 3 Accuracy, in percent, among the four classifiers

5. APPLICATIONS

A large portion of our work consisted of prototyping and testing potentially useful tools for the DJ. By sorting our dataset through the use of Self Organizing Maps, DJs will be able to generate groupings of musical material that immediately work well together. This data organization will provide not only obvious song clusters, but also interesting musical associations that may otherwise be overlooked.

5.1 Bricktable

A multi-touch surface called Bricktable [17] was chosen as the interface for visualizing and interacting with the SOMs. Multi-touch screens add a certain physicality to the data for the user, additionally supplying a modular software platform on which to expand the performance capabilities of these tools, especially between multiple potential users.

5.2 Self Organizing Maps

Our first application visualized the data in Processing using a SOM. The ability to effectively reduce dimensionality using a standard k-NN algorithm and the ease of visualization made a SOM an appealing choice to display the data as well as create a basic platform for playlist generation. The use of SOMs for playlist generation has been previously researched extensively by M. Dittenbach *et. al.* using their PlaySOM system [18].

Individual songs consist of 15 8-dimensional feature vectors. During our feature extraction stage, ChucK sends the features over Open Sound Control into Processing along with file name and path. The features are then ordered in a hierarchal manner, and superimposed over an RGB vector. The color vectors are then used to visualize unique songs on a 2D map. Once the map is populated and sorted, users can access individual songs by touching a coloured circle. This will recall the filename and begin playing the song, allowing users to quickly compare neighbouring music.

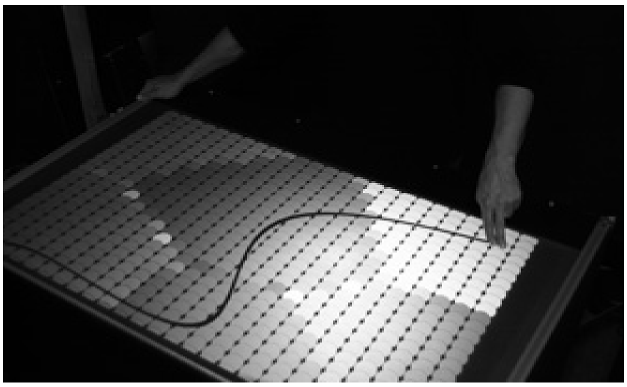

Figure 1 The SOM being displayed on the Bricktable, with the *DJen* interface minimized

5.3 The *DJen* Performance Application

Although the use of multi-touch for selecting songs directly on the SOM provides an engaging way for browsing a music collection, the application can be pushed further within the multi-touch paradigm. With this in mind, we present the *DJen* ('D-Gen') application, a tool to facilitate automatic set-list generation by enabling effective navigation of large libraries of music.

A critical skill among successful DJs is the ability to navigate seamlessly between many different songs, sometimes from varying genres. The key to this task is having the songs share a relationship in some way, usually through tempo. Via our SOM visualization, DJs already have access to musical groupings based off similarities, even if the genre is misclassified; however, *DJen* allows DJs to gesture a path through this map creating a dynamic playlist that can be used as source material for a performance. Due to the similarities between neighbours on the map, any arbitrary path will automatically generate a list of songs that share a strong relationship. Additionally, multiple DJs can create paths simultaneously, and *DJen* can interpolate a single path equidistant from all other paths. This will create a set list that represents the mean vectors between the original *DJen* paths. Finally, paths may be modified in real time for fine-tuning. Through this process we hope to enable the grouping of material in ways that a DJ may find inspiring. This path-based system is reminiscent of research conducted by R. Gulik and F. Vignoli in [19].

Figure 2 demonstrates the DJen GUI with the two primary UI elements shown: the playlist editor, and the 'now-playing' bar. Without a path set, a DJ can drag individual circles into the playlist editor to create a set list. When a path is drawn, the editor is automatically populated. If working collaboratively, another DJ may reshape the path and the playlist editor will automatically regenerate a set list.

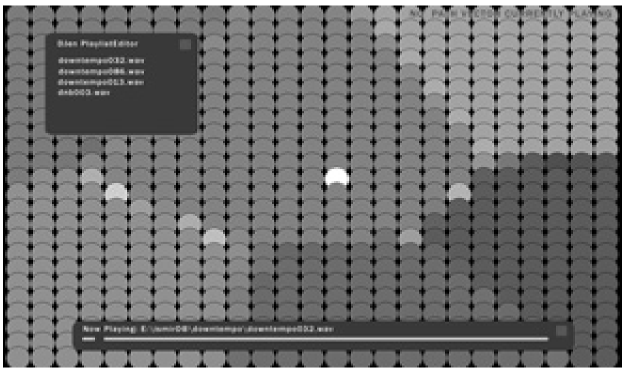

Figure 2 The SOM with the *DJen* GUI.

6. CONCLUSIONS & FUTURE WORK

Through the use of MIR tools we have shown strong potential for categorizing Electronica by genre. The k-NN algorithm provided an effective way of processing the sample data, making the *DJen* application possible through the use of a SOM. Finally, coupling this work with a multi-touch surface opened new avenues for DJs to interact with their music.

The *DJen* application represents a motivating step toward 'intelligent', MIR-powered tools for DJs. Its strength is revealed through the interactive user interface and visualization techniques. In the future, more rhythmic features to categorize electronica may be explored to create a system whereby *DJen* can perform automatic transitions between songs. This will allow the DJ to concentrate on other expressive areas such as sampling, looping, and effects processing.

The continuous growth of multi-touch necessitates development of further applications to explore both single and multi-user performance paradigms. Although *DJen* may be used by one or more users, extensive collaboration options may be enabled by allowing one DJ to oversee transitioning, while another manages multiple set-lists stemming from the original path chosen across the map.

Looking into the future, we hope *DJen* and other MIR-powered applications of its type will in the future enable any DJ to create expressive performances for their audiences.

7. ACKNOWLEDGEMENTS

We would like to thank Nick Diakopoulos for his helpful ideas and revisions. We would also like to thank George Tzanetakis for his inspiration and training.

8. REFERENCES

1. Gouyon, F. and S. Dixon. *Dance Music Classification: a tempo-based approach*. in *Proceedings of the 5th International Conference on Music Information Retrieval*. 2004. Barcelona, Spain.
2. Gouyon, F., et al. *Evaluating rhythmic descriptors for musical genre classification*. in *Proceedings of the 25th International AES Conference*. 2004. London, UK.
3. Tzanetakis, G., R. Jones, and K. McNally. *Stereo Panning Features for Classifying Recording Production Style*. in *Proceedings of the 8th International Conference on Music Information Retrieval*. 2007. Vienna, Austria.
4. Kapur, A., R. McWalter, and G. Tzanetakis. *Query-by-Beat-Boxing: Music Retrieval For The DJ*. in *Proceedings of the 5th International Conference on Music Information Retrieval*. 2004. Barcelona, Spain.
5. Birmingham, W., R. Dannenberg, and B. Pardo, *Query by humming with the VocalSearch system*. Commun. ACM, 2006. **49**(8): p. 49-52.
6. Kapur, A., R. McWalter, and G. Tzanetakis. *New Music Interfaces for Rhythm-Based Retrieval*. in *Proceedings of the 6th International Conference on Music Information Retrieval*. 2005. London, England.

7. Patten, J., B. Recht, and H. Ishii. *Interaction techniques for musical performance with table-top tangible interfaces*. in *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*. 2006. Hollywood, California.
8. Andersen, T.H. *Mixxx: Towards Novel DJ Interfaces*. in *Proceedings of the 2003 Conference on New Interfaces for Musical Expression*. 2003. Montreal, Canada.
9. Jorda, S., et al. *The reacTable*. in *Proceedings of the International Computer Music Conference*. 2004. Barcelona, Spain.
10. Hichner, S., J. Murdoch, and G. Tzanetakis. *Music Browsing Using a Tabletop Display*. in *Proceedings of the 8th International Conference on Music Information Retrieval*. 2007. Vienna, Austria.
11. Pampalk, E., G. Widmer, and A. Chan, *A new approach to hierarchical clustering and structuring of data with Self-Organizing Maps.* Intell. Data Anal., 2004. **8**(2): p. 131-149.
12. Cooper, M., et al., *Visualization in Audio-Based Music Information Retrieval.* Comput. Music Journal, 2006. **30**(2): p. 42-62.
13. Chen, Y.-X. and A. Butz, *Musicsim: integrating audio analysis and user feedback in an interactive music browsing ui*, in *Proceedings of the 13th International Conference on Intelligent User Interfaces*. 2009: Florida, USA.
14. Fiebrink, R., G. Wang, and P. Cook. *Support For MIR Prototyping and Real-Time Applications in the ChucK Programming Language*. in *Proceedings of the 9th International Conference on Music Information Retrieval*. 2008. Philadelphia, USA.
15. Witten, I.H. and E. Frank, *Data Mining: Practical machine learning tools and techniques*. 2nd ed. 2005, San Francisco: Morgan Kaufmann.
16. Duda, R.O., P.E. Hart, and D.G. Stork, *Pattern Classification*. 2nd ed. 2000, New York: John Wiley & Sons.
17. Hochenbaum, J. and O. Vallis. *Bricktable: A Musical Tangible Multi-Touch Interface*. in *Proceedings of the Berlin Open*. 2009. Germany.
18. Dittenbach, M., R. Neumayer, and A. Rauber. *PlaySOM: An Alternative Approach to Track Selection and Playlist Generation in Large Music Collections*. in *Proc. 1st Intl. Workshop on Audio-Visual Content and Information Visualization in Digital Libraries*. 2005. Cortona, Italy.
19. Guelik, R.V. and F. Vignoli. *Visual Playlist Generation on the Artist Map*. in *Proceedings of the 6th International Conference on Music Information Retrieval*. 2005. London, England.