Analysing and Evaluating Playlists on Music Maps

Jakob Frank

Vienna University of Technology, Vienna, Austria
http://www.ifs.tuwien.ac.at/mir
frank@ifs.tuwien.ac.at

Abstract. Paths drawn to a *Music Map* are used to quickly and intuitively create playlists by drawing figures onto the map surface. In this paper the correlation between the quality of a playlist and its visual shape on a *Music Map* is investigated. A series of amateur and professional playlists together with generated playlists are analysed according to their visualisation on a *Music Map*. Furthermore, the playlists are evaluated in a small user study where rated quality by the users is compared to the graphical representation and the song distances in different feature spaces.

1 Introduction

One question that arises whenever someone wants to listens to music is: "Is it worth creating a playlist – or do I just press *shuffle*?" The former is getting more and more difficult, due to the sheer amount of music available on today's computers while the latter is rendered useless by the huge variety of music on any player. One approach to ease the playlist creation process is to provide the user an intuitive and interactive overview of his music collection, a *Music Map*, and a quick way to generate playlists.

The quality of the generated playlists is, however, often not satisfying. To improve the quality of the generated playlists, it is necessary to understand what determines the quality of a playlist. Since *Music Maps* create playlists based on figures drawn on the map, a series of playlists were analysed according to their visual shape. Furthermore, a small user study was lauched to gain more information about the correlation of the qualitiy of a playlist and its visual shape on a *Music Map*.

The remainder of this paper is structured as follows: Section 2 gives a brief overview about the technical background of *Music Maps*. In Section 3 the approach to visualise playlists is presented. A small scale user study about the correlation of the visual shape of a playlist and its quality is presented in Section 4. Finally, Section 5 summarises the conclusions and gives an outlook to future work.

2 Technical Fundamentals

The analysis and visualisation of playlists as it is presented in this paper relies on *Music Maps* which provide a graphical interface to large audio collections. [1]

The creation of such a *Music Map* is divided into two steps: feature extraction and map creation. First, the individual songs are analysed to extract descriptive features from the content of the audio stream. A wide variety of different feature extraction algorithms is available. For the experiments in this paper, Rhythm Patterns were used. The feature extraction process for a Rhythm Pattern is composed of two stages. First, the spectrogram of the audio is computed using the short time Fast Fourier Transfomr (STFT). After that, the Bark scale, and other psycho acoustic modells are applied to the spectrogram, aggregating it to 24 frequency bands. The spectrogram is then transformed to the Sone scale which reflects the human loudness sensation. In the second step, a discrete Fourier transform is applied to the Sonogram, resulting in a (time-invariant) spectrum of loudness amplitude modulation per modulation frequency for each critical band. After further smoothing and weighting steps, Rhythm Patterns numerically represent magnitude of modulation for 60 modulation frequencies on 24 bands, thus resulting in a 1440-dimensional vector. High values for a specific modulation frequency in a number of adjacent bands indicate a specific rhythm occuring in a song. [3, 5]

After the feature extraction process, the resulting feature vectors are used as input for a self-organising map (SOM). A SOM is an neural network model that provides a projection from high dimensional data points to a lower, generally 2-dimensional output space. [2] A SOM iteratively arranges the data points in the output space is such a way, that data points which where located close to each other in the input space are also located nearby in the output space.

Applied to this domain, the SOM groups songs which share a common rhythm sturcture onto the same regions on the map. In detail, the algorithm works as follows. The map consists of a predefined number of units, which are arranged on a two-dimensional grid. Each of the units is assigned a randomly initialised model vector that has the same dimensionality as the input vectors. In each iteration, a randomly selected vector is matched with the closest model vector (winner). An adaptation of the model vector is performed by reducing the distance between the model vector and the feature vector. The neighbours of the winner are adapted as well, yet to a lesser degree than the model vector of the winning unit. Once the learning phase is completed, the feature vector of each music file is mapped to its best-matching unit on the map. The axes of the map have no specific meaning, the information is conveyed through the distances among the music files to each other.

One of the application scenarios of the resulting SOM is, in conjunction with different visualisations, as a *Music Map* giving an interactive and intuitive overview over a large audio collection. Furthermore, this allows to quickly create various playlists by drawing a path on the *Music Map* (see Figure 1). The system selects tracks that are located along the path, creating a sequence of songs following the "geographical" direction of the path. This results in a playlist re-

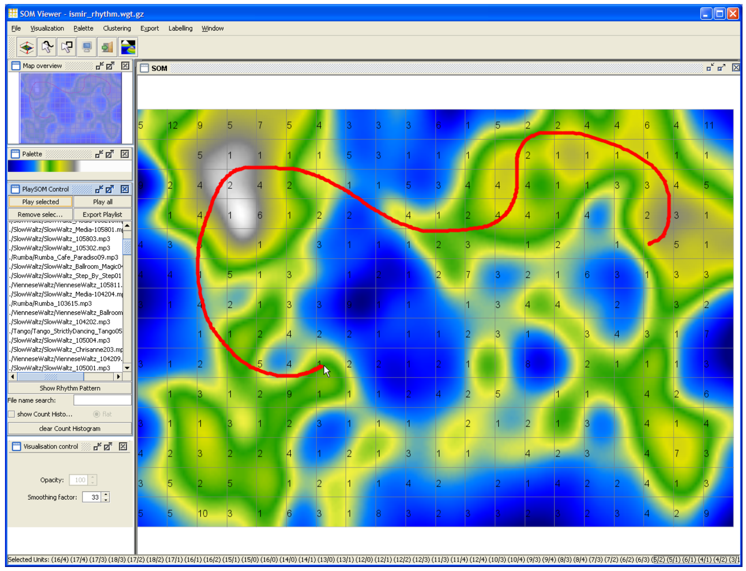


Fig. 1. PlaySOM showing a *Music Map* with a path and the corresponding playlist

stricted to a very specific musical style (which can, of course change over the progress of the playlist as the path moves on to some other region) but still containing a certain amount of variance. [4, 6]

3 Visualising Playlists

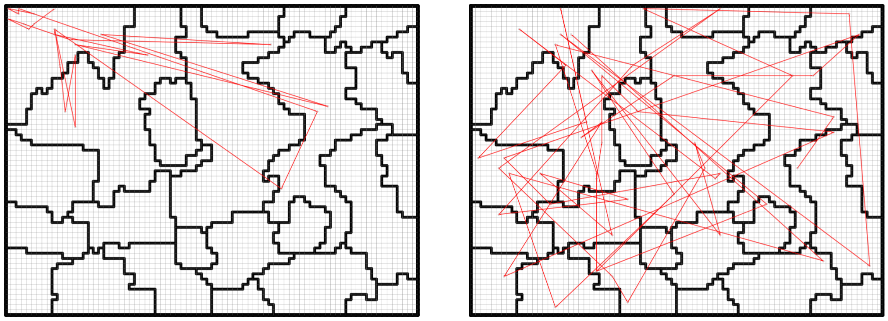
To visualise a playlist on the *Music Map*, for each song on the given playlist, the position on the map is located and linked with the positions of the adjacent songs oin the playlist. This creates a path based on a playlist. Such visualisations can be used as a template for new playlist (showing them as example or recommendation to the user), but they also reveal specific and descriptive information about the playlist. In conjunction with different visualisations of the *Music Map* the shape of the playlist and the regions it covers allows conclusions about the musical style and variance of the playlist.

3.1 Data Corpus

To analyse the visualisation of playlists, a data corpus has been created, based on playlists created by last.fm[1] users.

Last.fm users have the posibility to add songs they like to one of their personal playlist, which is stored together with their user profile. As users can view the profile of each other, it is also possible to view the playlists of a certain user.

[1] `http://last.fm`



(a) Homogenous and focussed playlist   (b) Playlist without clear focus

Fig. 2. Two playlists, visualised on a *Music Map*

Furthermore, it is also possible to listen to playlists that contain at least 45 tracks from at least 15 different artists[2].

The data corpus was created by randomly selecting approximately 900 last.fm users who had created at least one playlist with more than 20 songs. The playlists where fetched and saved, together with 30 second snippets of the of the songs which are available from last.fm. 30 second preview snippets are also available at amazon.com[3], so songs that could not be retrieved from last.fm were fetched from this alternative source.

This resulted in a data corpus of 31772 audio files and 960 playlists. A *Music Map* with 80×60 units was created for the visualisation and analysis of the playlists. For all experiments in this paper, only playlists with a coverage of 100% where used – reducing the size of the courpus of playlists from 960 to 82 playlists.

Finally, the corpus was further extended by adding playlists created by professional users. The song titles broadcasted by a popular mainstream pop radio station[4] where logged over several days and again 30 second snippets where fetched from last.fm or amazon.com. The playlists retrieved from the radio station where split at the full hour to avoid the gap created by news broadcast.

3.2 Analysis

The visual analysis of the playlists showed very heterogenious results. Some playlists clearly focussed onto a specific region on the map. Figure 2(a) shows an example of a playlist with a "good" shape, that stays in an area of the map with infrequent outliers with a larger distance.

[2] At the time of investigation, this feature was reserved for users willing to pay a monthly fee.

[3] `http://www.amazon.com`

[4] Oe3, `http://oe3.orf.at`

Figure 2(b) provides as an excellent counter-example. The playlist covers almost the entire map, jumping from one corner to the complete opposite. It further shows no clear focus for any region on the map.

Unfortunately, playlists of the latter shape are the clear mayority. Only a handfull of playlist formed a continuous path. Further investigations of the playlists and their paths showed, that playlists that focussed on specific regions in most cases only contained tracks from one artist, sometimes even only from one album, which is uncommon for most playlists, thus raising the questions whether the playlists retrieved from last.fm can be considered as “real” playlists, or if they are used as some kind of “bookmark list” for favourite songs, but, on the other hand, many reached from one end of the map to the other.

The visual appearance of the radio playlists was, however, about the same as the previous investigated playlists from last.fm. The main difference was that none of the radio playlists was homogenous enough to really focus on one specific area.

4 User Study

To further investigate the quality of the playlists from last.fm, and to determine whether the visual shape of a playlist allows conclusions about the quality of a playlist, a small scale user study was launched. Two playlists where chosen from the pool of available last.fm-playlists based on their graphical representation on the map, one with a “good” visual shape and one with a “bad” shape (covering a large part of the map, c.f. Figure 2). Further also two playlists from the radio station where picked analogous to the playlists from last.fm. Finally, two playlists generated through the map (by drawing a path onto the map) where included in the questionaire.

Before given to the participants, all playlists where truncated to equal length (of 13 songs) and their names and therefore their sources where concealed. For each of the six different playlists, the participants where asked to

- rate how good a song fits to its preceding song, on a scale from 1 to 5, where 5 represents the best value,
- give an overall rating of the playlist on the same scale,
- select up to three songs from the playlist that should be removed to improve the quality, and
- name situations and/or locations where this playlist would fit in.

4.1 Participants

The questionaire was completed by five participants (three male, two female). A user study in this size must not be considered as a representative evaluation, it is more a proof of concept for the questioning.

The participants where between 20 and 30 years old. All of them have attended natural science studies at a university; three of them already received a master’s degree, one is still a student and one droped out. Four of the participants are affiliated with a university or research institution.

Table 1. Average transition ratings by the participants. t is the average transition rating, o is the overall rating of the playlist.

Playlist	P-1 (m)		P-2 (m)		P-3 (f)		P-4 (m)		P-5 (f)		average	
Playlist	t	o	t	o	t	o	t	o	t	o	t	o
last.fm A	2.25	2	1.67	1	2.00	3	2.17	2	2.67	2	2.15	*2.0*
last.fm B	2.75	3	2.17	3	2.92	1	3.00	3	3.67	4	2.90	2.8
radio A	2.50	4	3.17	4	2.50	4	1.67	1	2.92	3	2.55	3.2
radio B	3.42	4	3.33	4	3.00	3	2.83	2	4.33	5	**3.38**	**3.6**
generated A	2.17	2	1.67	2	2.08	3	1.67	2	2.75	2	*2.07*	2.2
generated B	2.67	2	2.42	3	3.50	4	2.33	3	3.33	4	2.85	3.2

4.2 Results

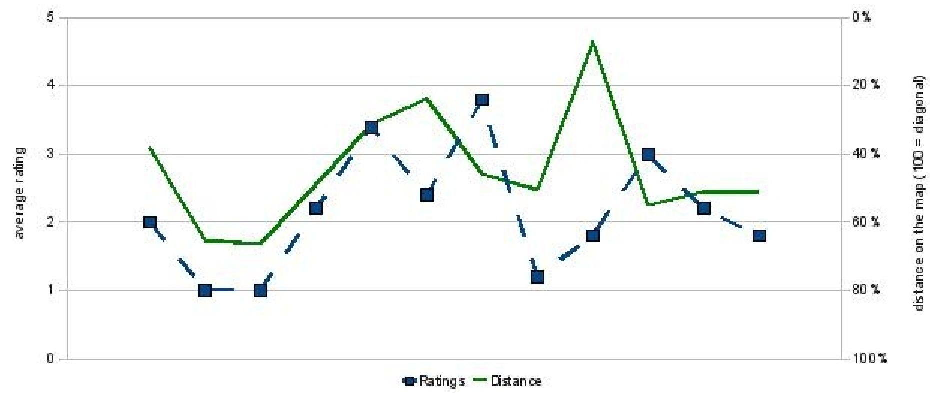
From the questionaires of the user study, summarised in Table 1, several interesting conclusions could be drawn:

1. Individual participants gave coherent answers. In most cases, the average rating of the transitions was about the same as the overall rating of the playlist.
2. The participants showed a clear preference for "professional" playlists taken from a radio station over all other playlists (c.f. Table 1).
3. Regarding the overall rating, generated playlists performed better than the playlists from last.fm. For the transition-rating, the playlists from last.fm where slightly in favour.
4. Regarding playlists from the same source, the one with the "better" visual shape also got higher ratings in all cases.
5. Whereas the individual rating of the transitions differ in most cases, the participants clearly agreed when naming the songs that would not fit into the playlist and thus should be removed.

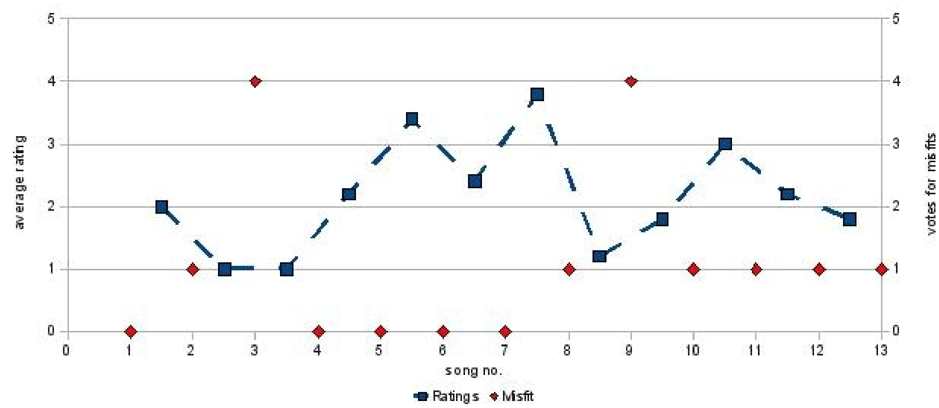
For all feature sets that have been investigated, no correlation between the users' rating and the distances in the feature space could be found. This observation can also be made for the distances on different maps trained with these feature sets.

This is illustrated in Figure 3 for a map based on Rhythm Patterns: in 3(a), the dotted line represents the average ratings by the participants (left Y-axis), while solid line represents the distances in the input space (right Y-axis, 100% is equivallent to the map's diagonal). The third song in the playlist has a notable longer distance to its neighbours in the list – these transitions also were rated rather low. The middle part of the playlist could be interpreted that closer distances got better ratings, but towards the end occurs a section where the closest distance in the playlist got very low ratings.

An interesting observation from Figure 3(b) ist the correlation between the ratings and the votes for misfitting tracks. In most cases, where the mayority of the participants found that a song does not fit into the playlist, also the transition to and from this song got notable lower ratings. Again, the playlist shown is exemplary, but the trend can be observed with all playlists.



(a) Partially correlation between the distance on the map and the users' ratings



(b) Transitions to and from outliers are rated low

Fig. 3. Graphical representation of one exemplary playlist *last.fm B*

5 Conclusion and Future Work

In this paper, the correlation between a playlist and its visual shape on a *Music Map* was investigated. It showed, that very homogenous playlists (by one artist or album) do stay very focussed in specific regions, but playlists with more diversity are distributed over large parts of the map. This is especially true for professional playlists. Future work on this part will be to investigate the behaviour of playlists on other maps based on a bigger, more heterogenious corpus.

The user study showed that the visual shape in the *Music Maps* used in this paper is not a sufficent quality measurement on its own, but it does allow conclusions regarding playlists from the same source. Furthermore it showed that also the feature vectors do not cover all aspects of playlist quality. It will be part of further investigations which additional aspects are required to sufficently describe the quality of playlists. Part of this work will comprise a refined user study with a more balanced and representative participants' list.

References

1. Jakob Frank. Enhancing music maps. In Frantisek Babic, Jan Paralic, and Andreas Rauber, editors, *Proceedings of the 8th International Student Workshop WDA 2008*, pages 53–59, Dedinky, Slovakia, June 26-29 2008.
2. Teuvo Kohonen. *Self-Organizing Maps*, volume 30 of *Springer Series in Information Sciences*. Springer, Berlin, Heidelberg, 1995.
3. Thomas Lidy and Andreas Rauber. Evaluation of feature extractors and psycho-acoustic transformations for music genre classification. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 34–41, London, UK, September 11-15 2005.
4. Robert Neumayer, Michael Dittenbach, and Andreas Rauber. PlaySOM and PocketSOMPlayer – alternative interfaces to large music collections. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 618–623, London, UK, September 11-15 2005.
5. Andreas Rauber, Elias Pampalk, and Dieter Merkl. Using psycho-acoustic models and self-organizing maps to create a hierarchical structuring of music by musical styles. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 71–80, Paris, France, October 13–17 2002.
6. Andreas Rauber, Elias Pampalk, and Dieter Merkl. The SOM-enhanced JukeBox: Organization and visualization of music collections based on perceptual models. *Journal of New Music Research*, 32(2):193–210, June 2003.