DESIGN AND EVALUATION OF A PROBABILISTIC MUSIC PROJECTION INTERFACE

Beatrix Vad,[1] Daniel Boland,[1] John Williamson,[1] Roderick Murray-Smith,[1] Peter Berg Steffensen[2]
[1]School of Computing Science, University of Glasgow, Glasgow, United Kingdom
[2]Syntonetic A/S, Copenhagen, Denmark
mail@bea-vad.de, daniel@dcs.gla.ac.uk, jhw@dcs.gla.ac.uk
roderick.murray-smith@glasgow.ac.uk, pbs@syntonetic.com

ABSTRACT

We describe the design and evaluation of a probabilistic interface for music exploration and casual playlist generation. Predicted subjective features, such as mood and genre, inferred from low-level audio features create a 34-dimensional feature space. We use a nonlinear dimensionality reduction algorithm to create 2D music maps of tracks, and augment these with visualisations of probabilistic mappings of selected features and their uncertainty.

We evaluated the system in a longitudinal trial in users' homes over several weeks. Users said they had fun with the interface and liked the casual nature of the playlist generation. Users preferred to generate playlists from a local neighbourhood of the map, rather than from a trajectory, using neighbourhood selection more than three times more often than path selection. Probabilistic highlighting of subjective features led to more focused exploration in mouse activity logs, and 6 of 8 users said they preferred the probabilistic highlighting mode.

1. INTRODUCTION

To perform information retrieval on music, we typically rely on either meta data or on 'intelligent' signal processing of the content. These approaches create huge feature vectors and as the feature space expands it becomes harder to interact with. A projection-based interface can provide an overview over the collection as a whole, while showing detailed information about individual items in context. Our aim is to build an interactive music exploration tool, which offers interaction at a range of levels of engagement, which can foster directed exploration of music spaces, casual selection and serendipitous playback. It should provide a consistent, understandable and salient layout of music in which users can learn music locations, select music and generate playlists. It should promote (re-)discovery of music and accommodate widely varying collections.

To address these goals we built and evaluated a system to interact with 2D music maps, based on dimensionally-reduced inferred subjective aspects such as mood and genre. This is achieved using a flexible pipeline of acoustic feature extraction, nonlinear dimensionality reduction and probabilistic feature mapping. The features are generated by the commercial Moodagent Profiling Service [1] for each song, computed automatically from low-level acoustic features, based on a machine-learning system which learns feature ratings from a small training set of human subjective classifications. These inferred features are uncertain. Subgenres of e.g. electronic music are hard for expert humans to distinguish, and even more so for an algorithm using low-level features [24]. This motivates representing the uncertainty of features in the interaction.

It is not straightforward to evaluate systems based on interacting with such high-dimensional data. This is not a pure visualisation task. Promoting understanding is secondary to offering a compelling user experience, where the user has a sense of control. How do we evaluate projections, especially if the user's success criterion is just to play something 'good enough' with minimal effort? We evaluated our system to answer:

1. Can a single interface enable casual, implicit and focused interaction for music retrieval?
2. Which interface features better enable people to navigate and explore large music collections?
3. Can users create viable mental models of a high-dimensional music space via a 2D map?

2. BACKGROUND

2.1 Arranging music collections on fixed dimensions

A music retrieval interface based on a 2D scatter plot with one axis ranging from slow to fast and the other from dark to bright on the timbre dimension is presented in [10]. The authors show this visualisation reduces time to select suitable tracks compared to a traditional list view. [11] presents a 2D display of music based on the established arousal-valence (AV) diagram of emotions [20], with AV judgements obtained from user ratings. An online exploration tool musicovery.com [6] enables users to select a mood

[1] http://www.moodagent.com/

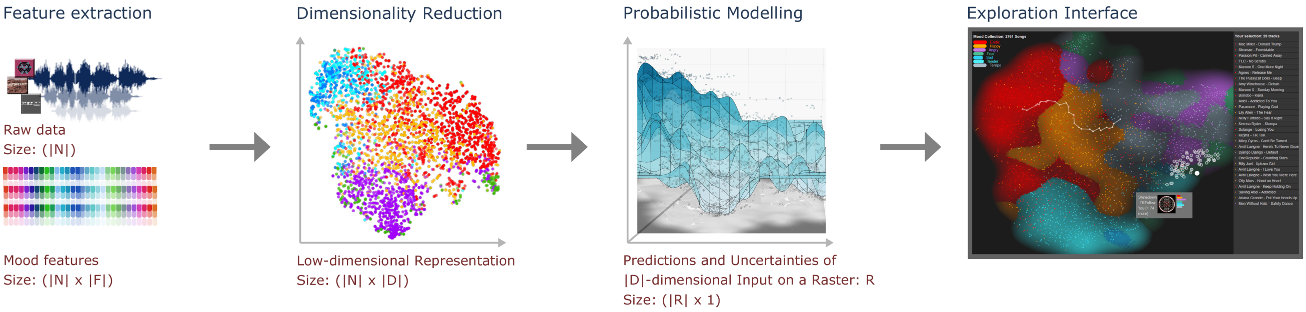

Figure 1. (a) An audio collection, described by a large set of features automatically extracted from the content. (b) visualisation of this high-dimensional dataset in two dimensions using dimensionality reduction (c) probabilistic models showing the distribution of specific features in the low dimensional space (d) combining dimensionality reduction with these models to build an interactive exploration interface.

in the AV space and starts a radio stream based on the input. These use two predefined dimensions that are easy to interpret, however they do not allow a broader interpretation of musical characteristics based on richer feature sets. [13] finds that music listening is often based upon mood. The investigation of musical preferences in [9] shows most private collections consist of a wide range of styles and approaches to categorisation.

2.2 Music visualisations via dimensionality reduction

"Islands of Music" [17] visualises music collections using a landscape metaphor. They use rhythmic patterns in a set of frequency bands to create a Self-Organizing Map (SOM), a map of music for users to explore. Similarly, [16] introduce the SOM-based PlaySOM and PocketSOM interfaces. Features are again based on rhythm and 2D embedding. An interesting visualisation feature is the use of "gradient fields" to illustrate the distribution of features over the map. Playlist generation is enabled with a rectangular marquee and path selection. Elevations are based on the density of songs in the locality, so clustered songs form islands with mountains. A collection of 359 pieces was used to evaluate the system and song similarities were subjectively evaluated. An immersive 3D environment for music exploration, again using a SOM is described in [14]. An addition to previous approaches is an integrated feedback loop that allows users to reposition songs, alter the terrain and position landmarks. The users' sense of similarity is modelled and the map gradually adapted. Both the SOM landscape and acoustic clues improved search times per song.

SongWords [2] is an interactive tabletop application to browse music based on lyrics. It combines a SOM with a zoomable user interface. The app is evaluated in a user study with personal music collections of ca. 1000 items. One reported issue was that only the item positions described the map's distribution of characteristics. Users had to infer the structure of the space from individual items. "Rush 2" explores interaction styles from manual to automatic [1]. They use similarity measures to create playlists automatically by selecting a seed song.

A detailed overview of music visualisation approaches and the MusicGalaxy system is contributed with [23]. This work introduces adaptive methods for music visualisation, allowing users to adjust weightings in the projection. It also explores the use of a lens so that users could zoom into parts of the music space. Most notably, it receives a significant amount of user evaluation. The lack of such evaluations in the field of MIR has been noted in [21], which calls for a user-centred approach to MIR. The work in this paper thus includes an 'in-the-wild' longitudinal evaluation, bringing HCI methodology to bear in MIR.

2.3 Interaction with music visualisations

Path drawings on a music visualisation, enabling high-level control over songs and progression of created playlists can be found in [26]. Casual interaction has recently started receiving attention from the HCI community [18], outlining how interactions can occur at varying levels of user engagement. A radio-like interface that adapts to user engagement is introduced by [3,4]. It allows users to interact with a stream of music at varying levels of control, from casual mood-setting to engaged music interaction. Music visualisations can also span engagement – from broad selections in an overview to specific zoomed-in selections.

3. PROBABILISTIC MUSIC INTERFACE

As shown in Figure 1, the interface builds on features derived from raw acoustic characteristics and transforms these into a mood-based visualisation, where nearby songs will have a similar subjective "feeling". Our feature extraction service provides over thirty predicted subjective features for each song including its mood, genre, style, vocals, instrument, beat, tempo, energy and other attributes. The features associated with moods chosen for later highlighting in the visualisation include *Happy, Angry, Sad* and *Tender*. These were identified as relevant moods from social tags in [12]. *Erotic, Fear* and *Tempo* (not strictly a mood) were also included. The features were investigated in [5].

Given our large number of features, we need dimensionality reduction to compress the data from $|F|$ dimensions to $|D|$ dimensions. The goal of this step is to preserve subjective similarities between songs and maintain coherent structure in the dataset. For interaction, we reduce down to 2D. We tried our system with a number of dimensionality reduction techniques including PCA and SOM. We chose the t-distributed stochastic neighbour embedding (t-SNE, [25]) model for non-linear dimensionality reduction to generate a map entangling a global overview of clusters of similar songs and yet locally minimise false positives.

To provide additional information about the composition of the low-dimensional space, we developed *probabilistic models* to visualise high dimensional features in the low-dimensional space. This probabilistic back-projection gives users insight into the structure of the layout, but also into the uncertainties associated with the classifications. On top of the pipeline (Figure 1), we built an efficient, scalable web-based UI which can handle music collections upwards of 20000 songs. The tracks can be seen as random variables drawn from a probabilistic distribution with respect to a specific feature. The distribution parameters can be estimated and used for prediction, allowing smoothed interpolation of features as shown in Figure 2. We used Gaussian Process (GP) priors [19], a powerful nonparametric Bayesian regression method. We applied a squared exponential covariance function on the 2D (x, y) coordinates, predicting the mood features P_f over the map. The GP can also infer the uncertainty σ_f^2 of the predicted feature relevance for each point [22].

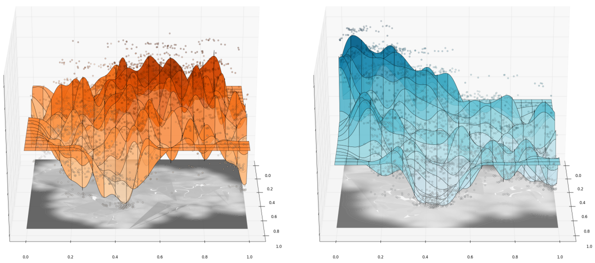

Figure 2. Gaussian Process predictions of features. Orange denotes the “happy” feature distribution and blue denotes “tender”. The greyscale surface shows the uncertainty; lighter is more certain and darker is less certain.

3.1 Interface design

To present the inferred subjective results to the users, the GP mean and standard deviation is evaluated over a 200 × 200 grid covering the 2D music space. A continuously coloured *background highlighting* is created where areas of high feature scores stand out above areas with higher uncertainty or lower scores. To highlight areas with high prediction scores and low uncertainty, a nonlinear transform is used: $\alpha_f = P_f^2 - \sigma_f^2$, for each mood feature f, having a standard deviation σ_f and a predicted feature value P_f. The clusters in the music space can be emphasised as in the upper part of Figure 3 by colouring areas with the colour associated with the highest score; i.e. $\text{argmax}(\alpha_f)$ – a winner-takes-all view. This not only divides the space into discrete mood areas but also shows nuanced gradients of mood influences within those areas. However, once a user starts to dynamically explore a specific area of the space, the system transitions to *implicit background highlighting* such that the background distribution of the mood with the highest value near the cursor is blended in dynamically as in the lower plots of Figure 3, giving the user more subtle insights into the nature of the space.

Tracks are represented as circles in a scatter plot, where size can convey additional information, e.g. the popularity of a song, without disturbing the shape paradigm. To support visual clustering, colour highlights the highest scoring mood feature of each song, and transparency conveys the feature score. However, the number of diverging, bright colours for categorisation is limited. Murch [15] states that a “refocus” is needed to perceive different pure colours, so matched pairs of bright and desaturated colours are chosen for the highly correlated mood pairs *tender*/*sad*, *happy*/*erotic* and *angry*/*fear*.

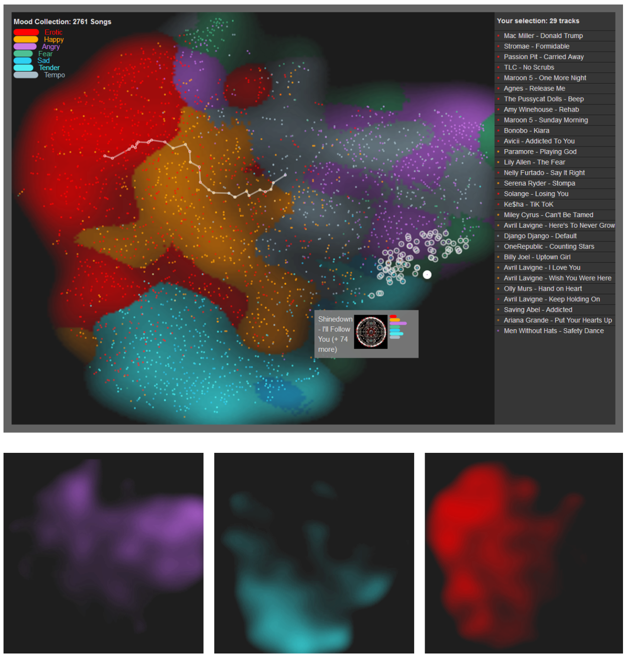

Figure 3. Top: The interactive web interface in its ‘winner takes all’ overview colouring. A path playlist selection as well as a neighbourhood selection is visible in the mood space. Bottom: Background highlighting for the features *angry*, *tender* and *erotic*. Compared with the overview colouring, the subtle fluctuations of features are apparent.

3.2 Interaction with the interface

As the visualisation can handle very large numbers of items, a *semantic zoom* was integrated, where the size of each element is fixed. This coalesces items on zoom out and disentangles items on zoom in.

Further insight into the nature of the space is given by the *adaptive area exploration* tool which visualises the local item density. In contrast to previous work we do not use a fixed selection area but one based on the k-nearest-neighbours to the mouse cursor. Points are highlighted as the mouse is moved, creating a dynamically expanding and collapsing highlight, responding to the structure of the space. The k-nn visualisation adapts to zoom level; when zoomed out, k is large; when zoomed in, we support focused interaction, with a smaller k.

Focus and context: To make the music map exploration more concrete, a hover box is displayed with information about the nearest item to the cursor, including artist name, title, and album art (see Figure 3). It shows a mini bar chart of the song's mood features. As this is fixed onscreen, users can explore and observe changes in the mood chart, giving them insight into the high-dimensional space.

3.3 Playlist generation

Neighbourhood selection is a quick and casual interaction metaphor for creating a playlist from the k nearest neighbours. Songs are ranked according to their query point distance. This enables the directed selection of clusters in the space, even if the cluster is asymmetric. By adjusting zoom level (and thus k), k-NN selection can include all in-cluster items while omitting items separated from the perceived cluster. This feature could be enhanced by adding an ϵ-environment similar to the density-based clustering algorithm DBSCAN [7]. Fast rendering and NN search was implemented using quadtree spatial indexing [8].

Path selection enables space-spanning selections. Drawing a path creates a playlist which 'sticks' to nearby items along the way. The local density of items is controlled by modulating velocity, so faster trajectory sections stick to fewer songs than slow ones. This 'dynamic attachment' offers control over the composition of playlists without visual clutter. E.g. a user can create a playlist starting in the *happy* area, then gradually migrating towards *tender*.

4. USER EVALUATION

The evaluation was based on the *research questions:*
1. How do users perceive the low-dimensional mood space projection? 2. Is the mood-based visualisation useful in music exploration and selection? 3. Which techniques do users develop to create playlists?

A pilot study evaluated the viability of the system and guided the design of the main longitudinal "in the wild" user study, which was conducted to extract detailed usage behaviour over the course of several weeks. Adapting to a new media interface involves understanding how personal preferences and personal media collections are represented. Longitudinal study is essential for capturing the behaviour that develops over time, beyond superficial aesthetic reactions and can – in contrast to Lab-based study – cover common use cases (choose tracks for a party, play something in the background while studying).

Eight participants (1 female, 7 male, 5 from the UK and 3 from Germany, undergraduate and research students) – each with their own Spotify account and personal music collections – were recruited. The mood interface was used to visualise the personal music collection of the participants. The participants used the interface at home as their music player to whatever extent and in whatever way they wanted. Two participants also used the system at work. All subjects used a desktop to access the interface. As a reward and to facilitate use together with the Spotify Web Player, participants were given a voucher for a three month premium subscription of Spotify.

The Shannon entropy H of the 6 mood features of each user's music collection gives an impression of the diversity of content. Using the maximum mood feature for each song, $H = -\sum_i p_i \log_2 p_i$, where $p_i = N_{mood_i}/N$.

	1	2	3	4	5	6	7	8
H	2.51	2.36	2.49	2.42	1.84	2.38	1.93	2.5
N	3679	2623	4218	3656	2738	2205	1577	3781

Table 1. Entropy H, no. tracks N of users' collections.

The study took place in two blocks, each with nominally four days of usage, although the actual duration varied slightly. One of the key aims was to find out if the probabilistic background highlighting provides an enhanced experience, so the study was comprised of two conditions in a counterbalanced within-subjects arrangement:
A Music Map without background highlighting.
B Music Map with background highlighting: The probabilistic models are included, with the composite view of the mood distribution as well as dynamic mood highlighting on cursor movements. Each participant was randomly assigned either condition **A** in week 1 followed by **B** in week 2 or vice versa. At the beginning of each condition and the end of the study, questionnaires were administered to capture participants' experience with the interface. Interface events, including playlist generation, navigation and all mouse events (incl. movements) were recorded.

5. RESULTS

Most participants used the software extensively, generating an average of 21 playlists per user per week, as shown in Table 2. On average, users actively interacted with the system for 77 minutes each week (roughly 20 minutes a day) – time spent passively listening to playlists is not included in this figure. Both groups generated more playlists in week 1 than in week 2, as they explored the system.

User	1	2	3	4	5	6	7	8
$N_{p,A}$	27	164	4	7	18	8	5	25
$N_{p,B}$	46	39	3	7	5	17	18	53

Table 2. No. playlists generated per user for cond. A & B. Users 1-5 had A in week 1, while 6-8 had A in week 2.

5.1 Mood perception

After each condition, users were asked to rate their satisfaction with interacting via the mood space. The overall opinion was encouraging. The majority of participants reported that they felt their collection was ordered in a meaningful way. Six stated that the mood-based categorisation made sense. Initially, the distinction of different music types was not rated as consistently over all conditions. This might be due to the fact that people usually discuss music in terms of genres rather than moods. However, the difficulty rating of mood changed over sessions. While six users rated mood-based categorisation as difficult at the start, only three participants still rated mood as difficult to categorise by the second week. This suggests that users can quickly learn the unfamiliar mood-based model.

5.2 Interactions with the Mood Space

Browsing the Space: Analysis of mouse movements provided insight into how participants explored mood spaces. Heatmaps were generated showing the accumulated mouse positions in each condition (Figure 4). Participants explored the space more thoroughly in week one of the study. Some participants concentrated exploration on small pockets, while others explored the whole space relatively evenly.

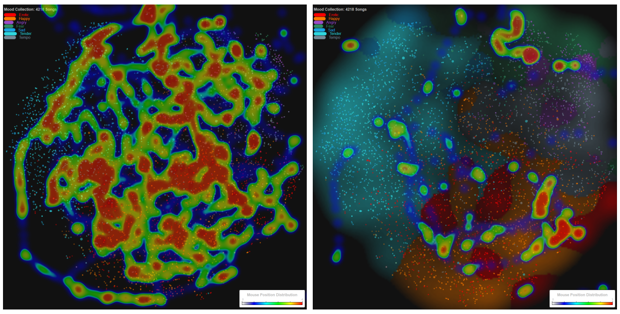

Figure 4. Heatmaps of interaction (mouse activity) of user 3 in week 1 (left) and week 2 (right). Interaction becomes more focused in the second week.

The browse/select ratio dropped noticeably for the second week for users with condition **A** first, as shown in Table 3. This suggests that participants browsed much more for each playlist in the first part of the study, and were more targeted in the second part. The browsing could have been either curiosity-driven exploration, or a frustrating experience, because the non-linear nature of the mapping made the space difficult for the users to predict the response to movements in the space. However, from Table 3 we can see that users who had the highlights in the first week seemed to have much more focused navigation from the start, and did not decrease their browsing much in week 2 when they lost the highlighting mechanism.

	Condition A	**Condition B**
Week 1	1218.39 *(483.49)*	447.94 *(176.85)*
Week 2	319.25 *(29.27)*	576.5 *(416.72)*

Table 3. Browse select ratios (std. dev. in parentheses) for week 1 and week 2 of the experiment, in cond. A and B.

Selections and Playlist Creation: Figure 5 shows playlists from two different participants in condition **B**. User 1 (left) created playlists by neighbourhood selection, and also drew a few trajectory playlists. User 7 (right) moved in over a more diverse set with a number of trajectory playlists. The paths partially follow the contours of the background highlight, which suggests this user explored contrasts in the mood space on these boundaries.

Neighbourhood selection was used more often (341 neighbourhood selections and 105 path selections). A rise in

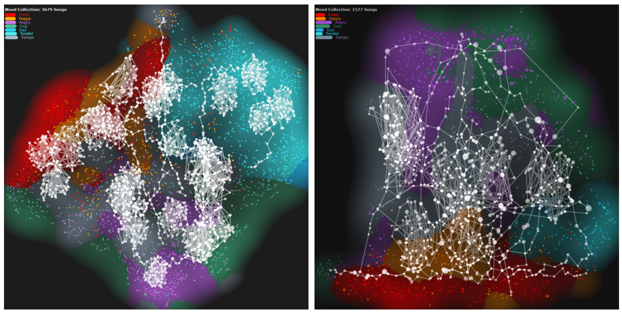

Figure 5. Created playlists under condition B for user 1 ($H = 2.51$) and user 7 ($H = 1.93$). Note the different class layouts for the collections with high/low entropy H.

the use of path selections, and a decline in neighbourhood selections can be seen in condition **B** versus **A**. In condition **A**, five times more neighbourhood than path selections were recorded, and only twice as many in condition **B** (see Table 4). This could be explained by the background distributions suggesting mood influence change gradually over the space. This information may encourage users to create trajectory playlists that gradually change from one mood intensity to another.

Selection	**A**	**B**	**Total**
Path	42	63	105
Neighbourhood	216	125	341
Neighbourhood/Path	5.1×	2.0×	3.3×

Table 4. Usage of the two different selection types in each condition. The neighbourhood/path ratio shows the increased use of the path tool in condition B.

5.3 Qualitative feedback

Background Highlighting: We asked whether background highlighting was valuable to the users. The answer was clearly in favour of background highlighting: 6/8 users valued the highlighting, one user was indifferent and one preferred the version without highlighting. The reasons given in favour of the highlighting were that they could more easily identify different regions and remember specific "locales" in the mood space. They recognised that songs had different mood influences and enjoyed following the colour highlights to areas of different intensity. One user stated that he liked the vividness of the implicit highlighting. The user who preferred no highlighting found it a cleaner look that was less confusing. 6 participants stated that they did not find the highlighting confusing. 7 participants answered that it did not distract from the playlist creation task. Qualitative feedback also indicated a preference for highlighting: *"[with highlighting] I could easier identify how the mood was distributed over my library", "coloured areas provided some kind of 'map' and 'landmarks' in the galaxy".*

Preference for neighbourhood versus path playlists: The domination of neighbourhood versus path playlists in the logged data is supported by feedback from questionnaires, which shows that users were generally happier with neighbourhood selection than the more novel path selection technique. The attitude towards the path selection differed, however, between conditions. Participants were more satisfied with path selection under condition **B**, with interactive background highlighting. After condition **A**, four participants agreed the path playlist was effective, and three disagreed. After condition **B**, however, 5 users agreed and only one disagreed.

Advantages of the Interface: The subjective feedback revealed that users had fun exploring the mood space and enjoyed the casual creation of playlists. *"fun to explore the galaxy", "easy generation of decent playlists"* Users also appreciated the casual nature of the interface: *"It was very easy to take a hands-off approach", "I didn't have to think about specific songs".* Users made specific observations indicating that they were engaged in the exploration task and learned the structure of the map, although this varied among users. *"I discovered that most older jazz pieces were clustered in the uppermost corner", "It was easy to memorize such findings [...] the galaxy thus became a more and more personal space".* Satisfaction with the quality of selections was high, although some participants found stray tracks that did not fit with neighbouring songs. *"The detected mood was a bit off for a few songs".* Several users stated that they appreciated the consistency of created playlists and the diversity of different artists, in contrast to their usual artist-based listening. There was concern that playlists did not offer enough diversity *"some songs that dominated the playlists", "too much weight given to 'archive' material", "some way to reorder the playlists to keep them fresh",* while others enjoyed this aspect: *"I rediscovered many songs I had not listened to in a long time".*

Shared versus personal: Visualising a shared (i.e. inter-user) mood space with personal collections embedded was not rated very important by most users (only one user thought this important). However, personalisation of the space was rated of high importance by half of the users. Ensuring that nearby songs are subjectively similar was additionally rated as important by the majority of participants (five users). These user priorities led to trade-offs between very large music maps and maps reliably uncovering intrinsic clusters of similar items.

Improvement requests: The most requested missing feature was a text search feature. The use of Spotify for playback also led to a disjointed user experience which would be easily improved on in a fully integrated mood-map music player. Users also requested the integration of recommendations and the ability to compare different mood spaces.

6. CONCLUSIONS AND FUTURE WORK

We presented an interactive tool for music exploration, with musical mood and genre inferred directly from tracks. It features probabilistic representations of multivariable predictions of subjective characteristics of the music to give users subtle, nuanced visualisations of the map. These explicitly represent the vagueness and overlap among features. The user-based, in-the-wild evaluation of this novel highlighting technique provided answers to the initial research questions:

Can users create viable mental models of the music space? The feedback from the 'in-the-wild' evaluation indicates that people enjoyed using these novel interfaces on their own collections, at home, and that mood-based categorisation can usefully describe personal collections, even if initially unfamiliar. Analysis of logged data revealed distinct strategies in experiencing the mood space. Some users explored diverse parts of the mood space and switched among them, while others quickly homed in on areas of interest and then concentrated on those. The questionnaire responses suggest they learned the composition of the space and used it more constructively in the later sessions. Users make plausible mental models of the visualisation – they know where the favourite songs are – and can use this model to discover music and formulate playlists.

Which interface features enable people to navigate and explore the music space? Interactive background highlighting seemed to reduce the need to browse intensively with the mouse (Table 3). Subjective feedback confirmed that it helped understand the music space with 6/8 users preferring it over no highlighting. Most users did not feel disturbed by the implicitly changing background highlighting. Both the neighbourhood and path playlist generators were used by the participants, although neighbourhood selections were subjectively preferred and were made three times more often than path selections. Subjective feedback highlights the contrast between interfaces which adapt to an individual user taste or reflect a global model, in which all users can collaborate, share and discuss music, trading greater relevance versus greater communicability. Similarly, how can we adapt individual user maps as the user's musical horizons are expanded via the exploratory interface? Users' preference of comparing visualisations over interacting in one large music space hints that an alignment of visualisations is a valid solution to this problem.

Can a single interface enable casual, implicit and focused interaction? Users valued the ability to vary the level of engagement. Their feedback also suggested that incorporating preview and control over the playing time of playlists would be useful, e.g. move towards "happy" over 35 minutes. A recurring theme was that playlists tended to be repetitive. One solution would be to allow the jittering of playlist trajectories and to do this jittering in high-dimensional space. The low-dimensional path then specifies a prior in the high-dimensional music space which can be perturbed to explore alternative expressions of that path.

Post-evaluation:
An enhanced version with a text search function was distributed at the end of the study. The encouraging result was that a month later, 3 of 8 participants still returned to the interface on a regular basis – once every few days, with one user generating 68 new playlists in the following weeks.

7. REFERENCES

[1] Dominikus Baur, Bernhard Hering, Sebastian Boring, and Andreas Butz. Who needs interaction anyway? Exploring mobile playlist creation from manual to automatic. *Proc. 16th Int. Conf. on Intelligent User Interfaces*, pages 291–294, 2011.

[2] Dominikus Baur, Bartholomaus Steinmayr, and Andreas Butz. SongWords: Exploring music collections through lyrics. In *Proc. ISMIR)*, pages 531–536, 2010.

[3] Daniel Boland, Ross McLachlan, and Roderick Murray-Smith. Inferring Music Selections for Casual Music Interaction. *EuroHCIR*, pages 15–18, 2013.

[4] Daniel Boland, Ross McLachlan, and Roderick Murray-Smith. Engaging with mobile music retrieval. In *MobileHCI 2015, Copenhagen*, 2015.

[5] Daniel Boland and Roderick Murray-Smith. Information-theoretic measures of music listening behaviour. In *Proc. ISMIR, Taipei*, 2014.

[6] Vincent Castaignet and Frederic Vavrille. (23.04.2014) http://musicovery.com/.

[7] Martin Ester, Hans P Kriegel, Jorg Sander, and Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Second Int. Conf. on Knowledge Discovery and Data Mining*, pages 226–231, 1996.

[8] Raphael A. Finkel and Jon Louis Bentley. Quad trees a data structure for retrieval on composite keys. *Acta Informatica*, 4(1):1–9, 1974.

[9] Alinka Greasley, Alexandra Lamont, and John Sloboda. Exploring Musical Preferences: An In-Depth Qualitative Study of Adults' Liking for Music in Their Personal Collections. *Qualitative Research in Psychology*, 10:402–427, 2013.

[10] Jiajun Zhu and Lie Lu. Perceptual Visualization of a Music Collection. In *2005 IEEE Int. Conf. on Multimedia and Expo*, pages 1058–1061. IEEE, 2005.

[11] JungHyun Kim, Seungjae Lee, SungMin Kim, and Won Young Yoo. Music mood classification model based on arousal-valence values. *13th Int. Conf. on Advanced Comm. Technology*, pages 292–295, 2011.

[12] Cyril Laurier, Mohamed Sordo, Joan Serrà, and Perfecto Herrera. Music mood representations from social tags. In *Proc. ISMIR*, pages 381–386, 2009.

[13] Adam J Lonsdale and Adrian C North. Why do we listen to music? a uses and gratifications analysis. *British Journal of Psychology*, 102(1):108–134, 2011.

[14] Matthias Lübbers, Dominik and Jarke. Adaptive Multimodal Exploration of Music Collections. In *Proc. ISMIR*, pages 195–200, 2009.

[15] Gerald M. Murch. Physiological principles for the effective use of color. *Computer Graphics and Applications, IEEE*, 4(11):48–55, 1984.

[16] Robert Neumayer, Michael Dittenbach, and Andreas Rauber. PlaySOM and PocketSOMPlayer, alternative interfaces to large music collections. In *Proc. ISMIR*, pages 618–623, 2005.

[17] Elias Pampalk, Andreas Rauber, and Dieter Merkl. Content-based organization and visualization of music archives. *Proc. 10th ACM Int. Conf. on Multimedia*, page 570, 2002.

[18] Henning Pohl and Roderick Murray-Smith. Focused and casual interactions: allowing users to vary their level of engagement. In *Proc. ACM SIGCHI Conf. on Human Factors in Computing Systems*, pages 2223–2232, 2013.

[19] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[20] James A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–1178, 1980.

[21] Markus Schedl and Arthur Flexer. Putting the User in the Center of Music Information Retrieval. In *Proc. ISMIR*, Porto, Portugal, 2012.

[22] Devinderjit Sivia and John Skilling. Data Analysis: A Bayesian Tutorial. *Technometrics*, 40(2):155, 1998.

[23] Sebastian Stober. *Adaptive Methods for User-Centered Organization of Music Collections*. PhD thesis, Otto-von-Guericke-Universitt Magdeburg, 2011.

[24] Bob L. Sturm. A simple method to determine if a music information retrieval system is a horse. *IEEE Transactions on Multimedia*, 16(6):1636–1644, 2014.

[25] Laurens van der Maaten and Geoffrey Hinton. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.

[26] Rob van Gulik and Fabio Vignoli. Visual playlist generation on the artist map. In *Proc. ISMIR*, pages 520–523, 2005.

Acknowledgments This was supported in part by the Danish Council for Strategic Research of DASTI under the CoSound project, 11-115328