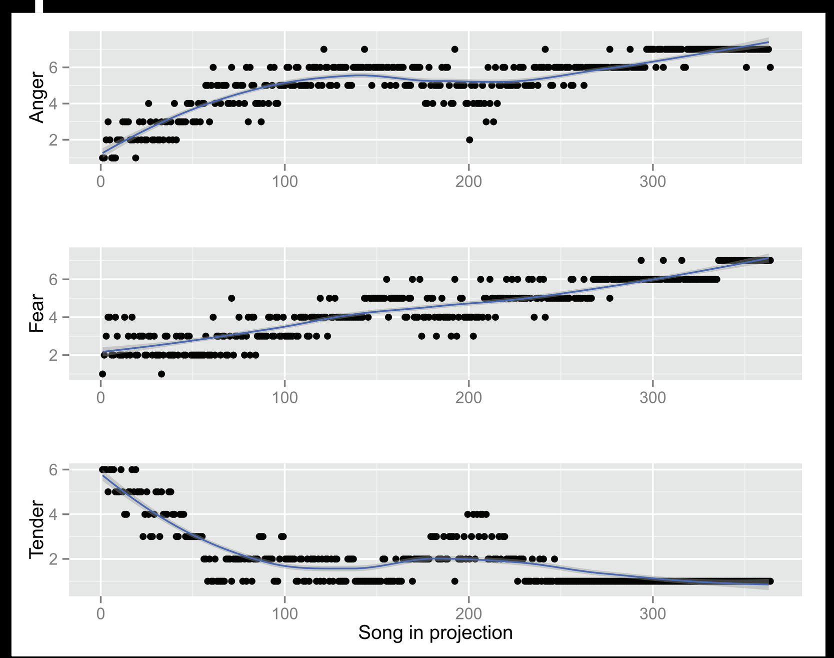
Finetuner: *Inferring Intended Music in a Shared Control Radio Interaction*

Daniel Boland, Ross McLachlan, Roderick Murray-Smith
School of Computing Science, University of Glasgow
Contact: daniel@dcs.gla.ac.uk

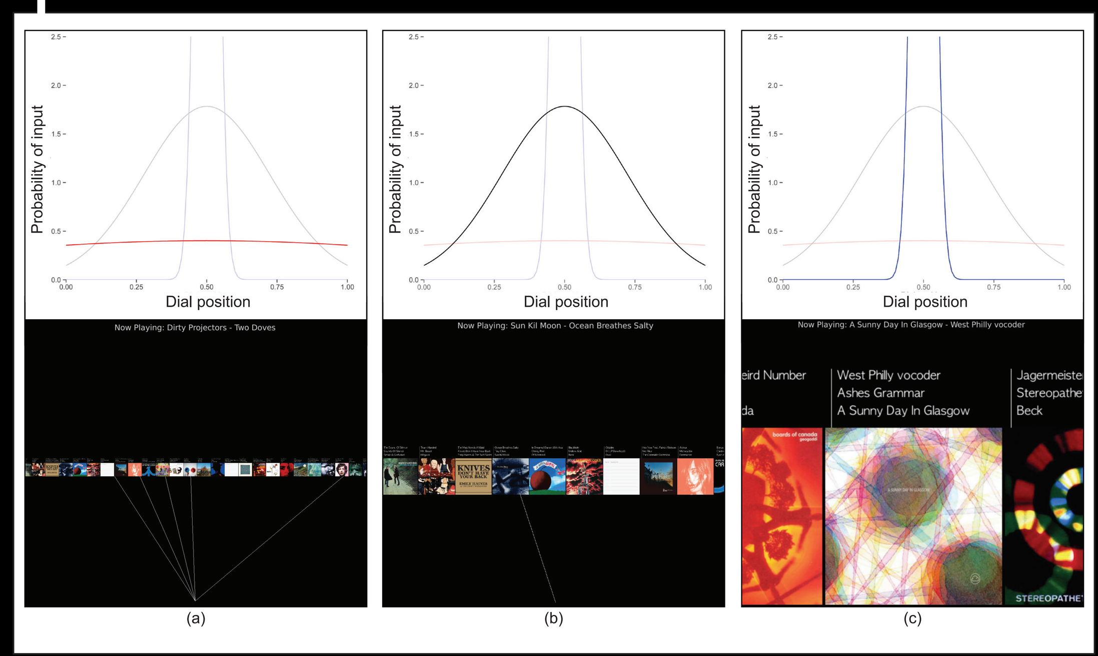
Projecting the Music Space

Six mood features are acquired for the user's music collection using MoodAgent[1]. This mood space is projected down to one dimension, using the NeRV[2] algorithm. This provides a way of arranging the music collection by mood to allow it to be navigated with a radio-like interaction. As can be seen in the figure below, the mood of the music slowly changes through the collection. In this collection, 'fear' increases steadily, as does anger in general however with an area of less angry, tender and fearful music.



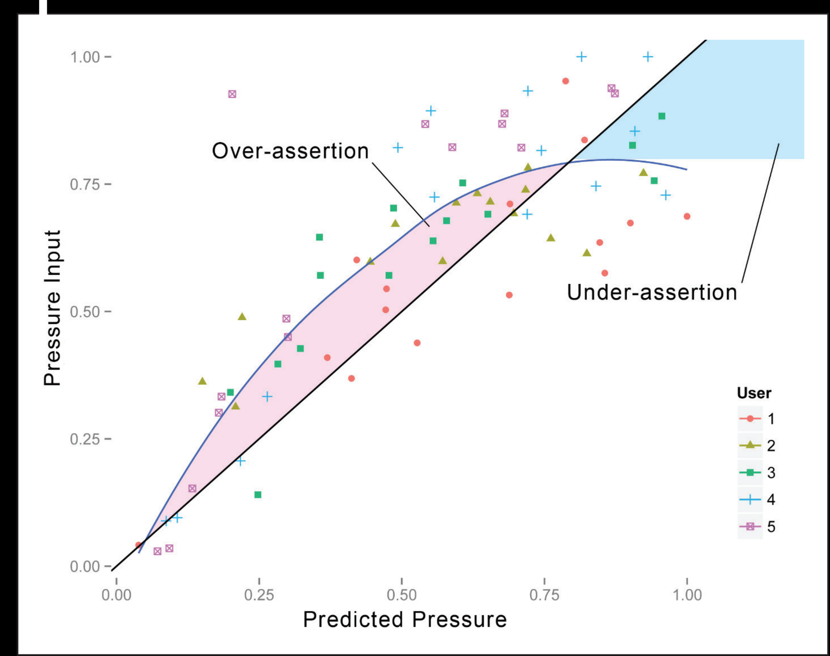
Predicting Input & Inferrring a Selection

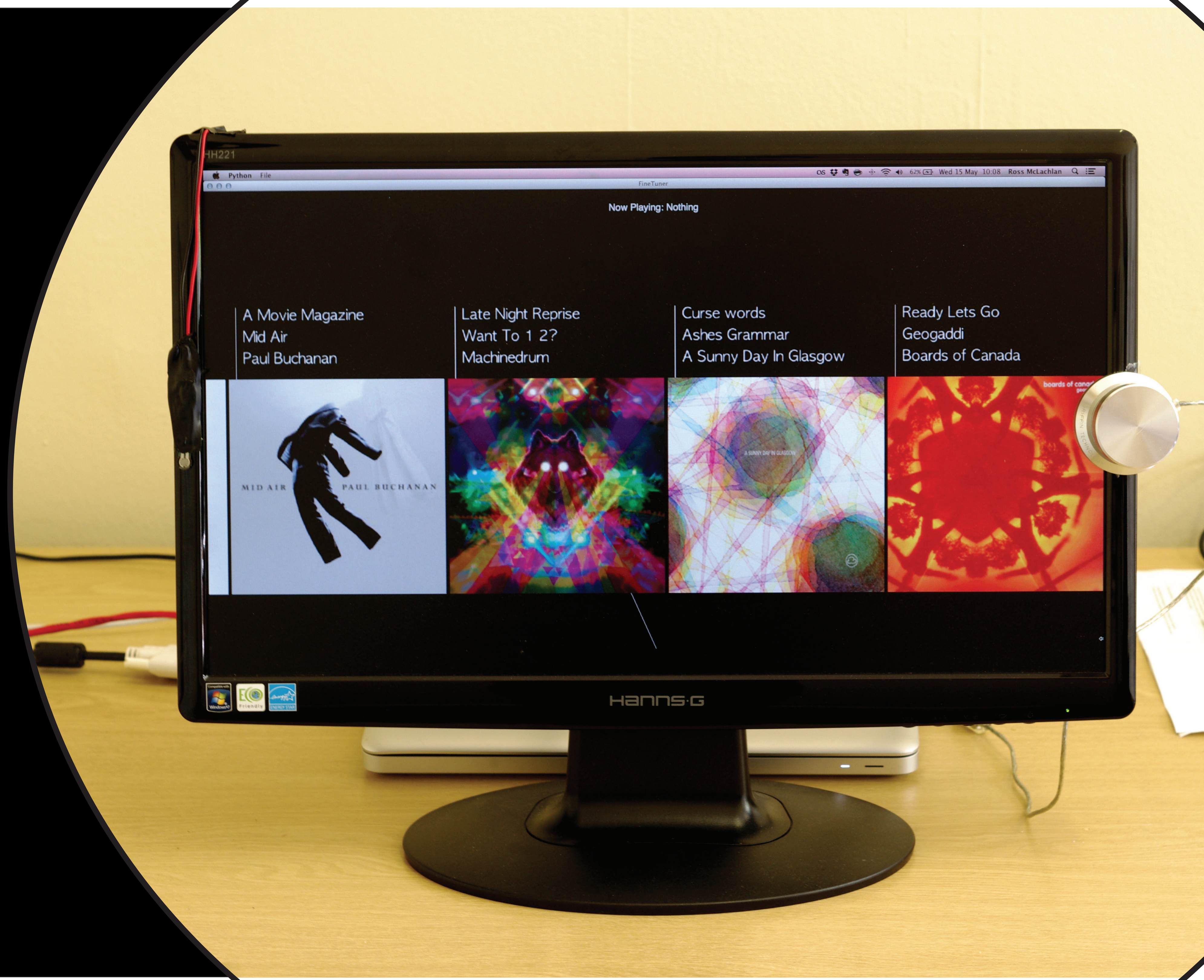
Users browse the music using a dial and can assert more control by applying pressure to a pressure sensor.[3] The system predicts the user's input for a given level of assertion, for example:
(a) The user asserts little control, a wide range of inputs is predicted, a broad selection is inferred using evidence from listening history.
(b) The user is applying more control and so the distribution of inputs predicted for a given song is narrower. A more explicit selection is inferred, with less weight given to previous listening history.
(c) At full assertion, the system predicts a precise input for each of the albums, resulting in specific album selection.



Iterative Design & Evaluation

As well as evaluating the system by running design sessions with users, the use of a generative model to predict user input gives an additional evaluation technique. Random fixed selections were shown to users who then had to provide the input they expect would produce such a selection. The figure below shows how well the user input matched the predicted input. Whilst users broadly understood what input to provide, the under and over assertion reigons highlight where the model can be improved.





Outline

We present a novel radio-like music interaction, featuring:

- Shared control between system and user over music selection
- Pressure sensor for users to assert control over the system
- Dial which can be turned to scroll through the music space
- Generative model of user input to enable inference of intended songs, incorporating prior knowledge from listening history.
- Visualisation of the posterior distribution over the music space to give feedback about uncertainty in the input.
- User-centered approach to inference, iterating our generative model using user data.

The generative model can be used as a likelihood function for user input. This likelihood function enables the use of Bayesian Inference to infer selections from user input - similar to query likelihood models in Information Retrieval.[4]

We model user input as being sampled from a Gaussian distribution around the *x* position of a song in the music projection:

$$p\left(i_x\right) = \sqrt{\frac{\tau}{2\pi}}\, e^{\frac{-\tau(i_x - x_s)^2}{2}}$$

This width of this distribution is controlled using a precision parameter τ which is proportional to the asserted control (pressure input):

$$\tau = \frac{A}{\sigma_d^2}$$

The belief about a song s_i being of interest to the user for a given input i_x is inferred using the input likelihood model and prior belief $p(s_i \,|\, A)$ over the music space obtained using LastFM music listening data:

$$p\left(s_i \,|\, i_x, A\right) = \frac{p\left(i_x \,|\, s_i, A\right) p\left(s_i \,|\, A\right)}{p\left(i_x \,|\, A\right)}$$

User Study

A user study was conducted with three particpants, with fully personalised music spaces generated using their music listening history and album art from their Last.fm accounts:

- All participants use shuffle and explicit selections of music.
- Participants felt shuffle favoured highly played songs too much.
- Participants expressed dislike when dissimilar songs appeared next to each other.
- Two suggested that attending concerts is relevant evidence for inferring what music to play.

Participants enjoyed the interaction technique, highlighting its advantages over shuffle or menu-based interfaces. Some areas for improvement were identified such as tuning the projection to keep dissimilar items apart and using additional evidence e.g. concert attendance.

References

[1] MoodAgent. Syntonetic (2013). *http://www.moodagent.com*

[2] Venna, J., Peltonen, J., Nybo, K., Aidos, H., and Kaski, S. Information Retrieval Perspective to Nonlinear Dimensionality Reduction for Data Visualization. *Journal of Machine Learning Research 11* (2010), 451–490.

[3] Interlink FSR Integration Guide and Evaluation Parts Catalog Page with Suggested Electrical Interfaces.

[4] Lafferty, J. and Zhai, C. Document language models, query models, and risk minimization for information retrieval. *In Proc. SIGIR*, ACM (2001), 111-119.

University of Glasgow | School of Computing Science