Adapting computational music similarity models to geographic user groups

Daniel Wolff

Department of Computer Science, MIRG, City University London, UK
daniel.wolff.1@city.ac.uk – http://juppiemusic.com

Tillman Weyde

Department of Computer Science, MIRG, City University London, UK

We present first results of experiments using music similarity ratings from human participants for group-specific similarity prediction. Music similarity is a key topic of research in music psychology and ethnomusicology. Computational models of music similarity have many applications such as music recommendation and indexing of music databases.

This study evaluates the feasibility of adapting similarity models to location-specific subsets of similarity ratings. To this end we use information on the country where the data was provided. Apart from directly training similarity models to the localised data, we perform a gradual adaptation of a previously trained general similarity model to the location-specific data. This allows us to compare the general and localised similarity models, providing a comparative analysis of the importance of acoustic features (e.g. loudness, timbre, tempo, chroma, key) for modelling similarity judgment across user groups. In future work, such groups could be selected to yield culturally determined models.

Our results show that localised models can be trained, but in comparison to general models this task proves more difficult due to the relatively small amount of training data available from each country. We found that the performance for some localised models can be increased using a general model as a basis for training. In one case this allows for the analysis of relevance of individual features for the specific data.

The similarity ratings used in our experiments were collected in the online Game With A Purpose "Spot The Odd Song Out". The mostly popular music presented in the game is based on the openly available MagnaTagATune and Million Song datasets, two large music datasets that come with acoustic descriptors for the music. Additionally to the similarity data being collected via triad questions, the modular game architecture allows for the collection of other human annotations, such as timbre and rhythm data. We also describe the extensible game with a discussion of further possibilities of its application.