Sound Similarity: Adding a New Dimension to Music Discovery
Franz de Leon, Kirk Martinez, Robert Damper
Electronics and Computer Science

Background
“Digital downloads driving recorded music industry towards recovery” [1]

- Global recorded music revenues up 0.3%, boosted by downloads, subscription and other channels
- Digital revenues up 9%, with major music services now open in more than 100 markets
- Music is helping fuel the digital economy, but barriers to growth need to be addressed

With the growing appetite for digital music, there is a need for new applications for browsing, organising, discovering and generating playlists.

Searching similar music

- Collaborative filtering
- Social tag data
- Expert metadata
- Direct analysis of audio content

Research objectives

The research aims to investigate how content-based methods can be used to perform music similarity estimation. This work hopes to complement music search engines to provide an exciting journey of music discovery.

Key issues

- What features are essential?
- Why are these features essential?
- How are these features extracted?
- How are these features summarized?
- How are the distances between features computed?
- What statistics are used to evaluate the algorithm?

Initial results

Precision is the ratio of the songs retrieved that are similar to the query song. The plot below shows the results for a database [5] with 1000 songs from 10 genres, 100 songs per genre. The system works best on classical, pop and metal. Moreover, the proposed system takes only 15 msec. to retrieve tracks.

Feature extraction

Timbre is “that attribute of auditory sensation in terms of which a listener can judge that two sounds, similarly presented and having the same loudness and pitch, are dissimilar.” [2] Hence, it is crucial to develop a computational model that captures the salient features of timbre.

Retrieval

Music retrieval is done in two stages: broad search and fine search. Broad search quickly generates a list of candidate similar songs to a query song using a simple Euclidean distance search. Fine search performs on the candidate songs using SKL divergence [3].

Acknowledgements

Franz de Leon is supported by the Engineering Research and Development for Technology Faculty Development Program of the University of the Philippines, and DOST T echnology Faculty Development Program of the University of the Philippines, and DOST.

References