Dublin Institute of Technology
ARROW@DIT

Conference papers

Audio Research Group

2008-01-01

Interacting with Large Music Collections: Towards the use of Environmental Metadata

Gordon Reynolds
Dublin Institute of Technology

Dan Barry
Dublin Institute of Technology, dan.barry@dit.ie

Ted Burke
Dublin Institute of Technology, ted.burke@dit.ie

Eugene Coyle
Dublin Institute of Technology, Eugene.Coyle@dit.ie

Follow this and additional works at: http://arrow.dit.ie/argcon

Part of the Electrical and Computer Engineering Commons

Recommended Citation
Raynolds, G., Barry, D., Burke, T., Coyle, E.: Interacting with large music collections:towards the use of environmental metadata. IEEE International Conference on Multimedia and Expo, Hanover, June 23-26, 2008, pp.989-992. Publication Date: June 23 2008-April 26 2008.

This Conference Paper is brought to you for free and open access by the Audio Research Group at ARROW@DIT. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@DIT. For more information, please contact yvonne.desmond@dit.ie, arrow.admin@dit.ie.

INTERACTING WITH LARGE MUSIC COLLECTIONS: TOWARDS THE USE OF ENVIRONMENTAL METADATA

Gordon Reynolds; Dan Barry
Ted Burke; Eugene Coyle

{gordon.reynolds; dan.barry; ted.burke; eugene.coyle}@dit.ie

Audio Research Group, Electrical Engineering Systems,
Dublin Institute of Technology, Ireland

ABSTRACT

Large music collections afford the listener flexibility in the form of choice, which enables the listener to choose the appropriate piece of music to enhance or complement their listening scenario on-demand. However, bundled with such a large music collection is the demanding task of manually searching through each entry in the collection to find the appropriate song required by the listener. This paper highlights the need for contextual and environmental information, which ultimately defines the listener's listening scenario.

Here, the preliminary results of an online music survey are analysed. These results indicate the possibility of how environmental features may be used as metadata to indicate the listener's mood. Therefore, environmental features, such as location, activity, temperature, lighting and weather have great potential as metadata and hence may be used to create a personalised automatic playlist generator for large music collections.

***Index Terms*—** Large Music Collections, Automatic Playlist Generation, Environmental Metadata, Mood Detection, Music Retrieval

1. INTRODUCTION

Technology is often held responsible and associated with creating the semantic gap. As outlined by Van Noorden [1], modern large music collections demonstrate this principle. Due to the efficiency of the modern audio codec, it is possible to maintain quality while minimising the storage space required for each song. In addition to this, advances in storage technology afford the user the possibility of storing thousands of songs on pocket sized music players. It is also estimated, that the next generation of MP3 players will provide listeners with access to millions of songs on a single device [2]. This implies that these devices could store up to 150,000 times the amount of songs when compared to current models [2]. In this situation, technology has created the semantic gap of music access. Current portable music players do not allow ease of access to large music collections.

This paper commences with a brief overview on how the problem of accessing a large music collection has been tackled in the past. The paper then continues to discuss metadata, in particular how the current available metadata may be expanded to include environmental metadata to represent the listener's mood in the automatic song selection process.

2. ACCESSING LARGE MUSIC COLLECTIONS

The problem of accessing large music collections has been defined and implemented in several ways. Such solutions have given rise to both simple and complex algorithms for automatically generating music playlists [3] [4], the simplest being 'shuffle play'. To generate meaningful automatic playlists, the system requires song information, or tags, often referred to as metadata.

2.1. Sorting Music Collections

Metadata is often used to index or sort a music collection, as well as retrieve music from a collection. Metadata is loosely defined as data about data [5]. The AES has described two primary forms of metadata [6]: 1) Manual Metadata and 2) Automatic Metadata. These are indicated in Figure 1.

An example of manual metadata includes the ID3 tags which provides the listener with artist and album information, amongst other things. Automatic metadata

includes context and sensor data, such as the listener's playback history.

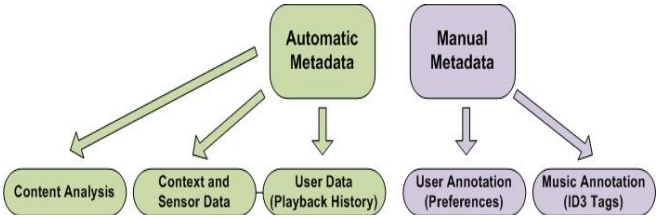

Figure 1: A characterisation of typical metadata for music [6].

2.2. Music Selection Using Metadata

With the availability of such rich metadata, there are several different systems available for generating automatic playlists. These include the use of a '*seed song'* or the use of a process known as '*collaborative filtering'*.

2.2.1. Playlist Generation Using Seed-Songs.
Within the 'seed song' approach, playlists are generated by providing the system with a song, the 'seed song', after which the system automatically generates a playlist of songs which '*sound*' similar to the seed song [3]. In general however, such systems tend to create uninteresting playlists that have little musical variation.

2.2.2. Playlist Generation Using Collaborative Filtering
Playlists may also be generated using a technique known as collaborative filtering [4]. Collaborative filtering is a community and voting process in which listeners with similar musical tastes are grouped together and share each others music collection. In contrast to providing a system with a *'seed song'*, collaborative filtering may provide a more varied and interesting playlist.

In either case, each technique uses a combination of automatic and manual metadata. In particular, *content* analysis is implemented using a sophisticated array of DSP algorithms to extend the metadata associated with each song, such as Lieue's comb-filter technique to track a song 's tempo [7]. Manual metadata is integrated into these systems with the use of ID3 tags which provide information such as song name, artist and album. Figure 2 is an example of two feature spaces associated with each song, namely a textual and a music feature space. However, any number of dimensions and feature combinations can potentially be used

The selection process is predominantly ruled by the emotional state and attitude of the individual [8]. That is to say, individuals are a function of their mood [9] and music selection is no different. Therefore, to provide a listener with a meaningful personalised automatic playlist generation system, it is ideal for the system to consider the listener's mood.

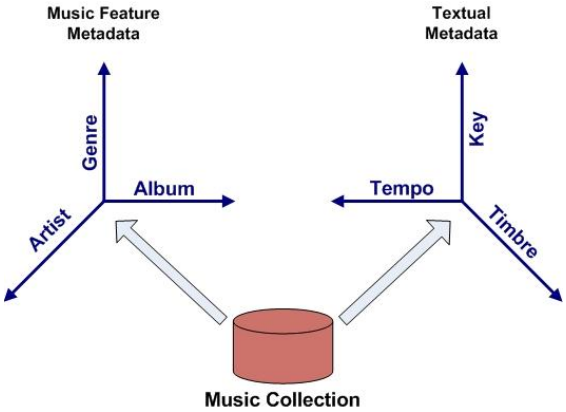

Figure 2: Shows the textual and musical feature space of a music collection.

3. DERIVING MEANING FROM MUSIC

As described by McDonald [10], primarily, meaning can be derived from music in three ways. These are 1) Cultural, 2) Social and 3) Association. These are indicated in Figure 3.

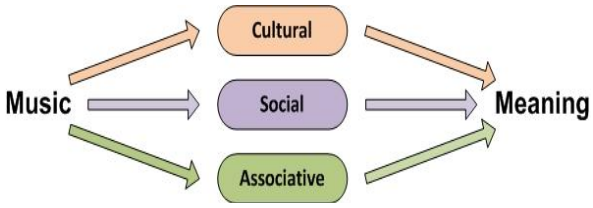

Figure 3: Outlines the three main processes in how people derive meaning from music [10].

In general, both the cultural and social meaning of music is a group or community influence interpretation of music rather than individual [10]. Such an example of this includes film scores, where the whole audience may feel a similar heightened sense of emotion within a movie scene through the interpretation of music [10].

When considering these three attributes of meaning, in reference to personalised selection, association provides the largest challenge to over come. This is because music association is the most personal and individual of the three attributes [10]. An individual may associate particular songs with a multitude of emotions and life situations.

As a result, this may invalidate the process of meta-tagging each song with an emotion directly, as the emotion-tags are not transportable between listeners, as described by *Tolos et al.* in [11]. However, an alternative approach is considered where the environment may be used to represent

emotions and hence generate a meaningful automatic playlist.

3.1. Using the Environment to Represent Emotions

Since measuring a listener's mood, or emotional state, directly boarders on impossibility, a mood indicator is required that is simple to measure and exploit. Hence, the listener's environment is suggested.

With the establishment of attitude theory, strong links have been forged between an individual's environment and their attitude, which in turn defines their mood and behavior [8]. The experience of an individual imposed by the outside world reflects how they feel on the inside [9]. Hence, the environment, and therefore environmental metadata, may be used as a mood indicator. Such a process is described in [12].

Figure 4 shows the integration of environmental metadata into the already existing song metadata spaces, namely a textual and a music feature space.

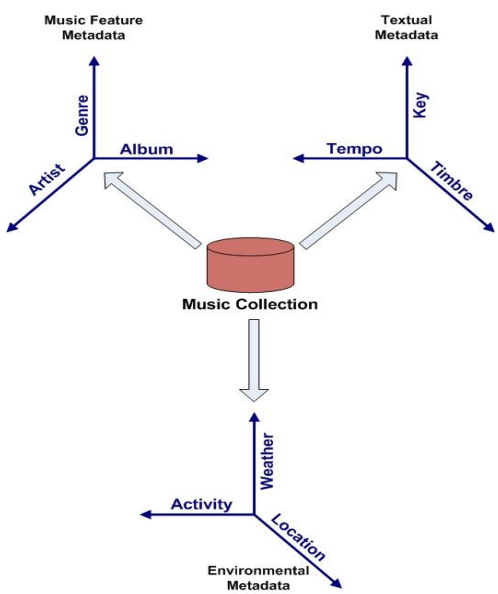

Figure 4: Shows the textual, musical feature and environmental space of a music collection

5. RESULTS

This section examines preliminary results from a survey recently taken in which over 750 music enthusiasts participated. This survey can be found online and completed at the following URL, www.audioresearchgroup.com/survey. The survey examines how listeners listen to music as well as examining how external influences can affect the listener's choice of music and mood.

To commence, it is noted that there is a relationship between the numbers of songs a listener has in their collection to the type of technology used, Figure 5.

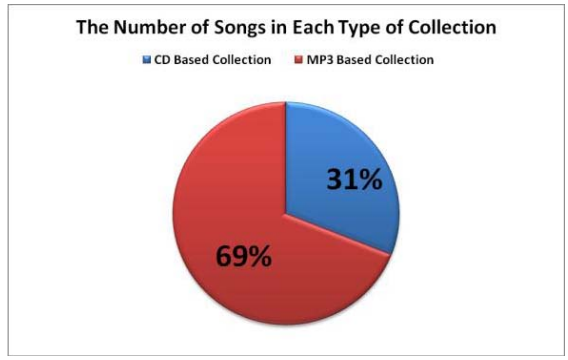

Figure 5: The number of songs in a music collection compared to the type of technology used.

As indicated by Figure 5 listeners that own an MP3 collection have more than double (69%) the number of songs that a listener has on CD (31%). This result supports the earlier discussion that modern audio technology, storage and interconnectivity allows the listener to store and transport more music.

In addition, results have been analysed for the affect that the environment has on a listener's mood and a listener's music selection. Initial environmental features monitored included temperature, weather, activity, lighting and location.

As all parameters can not be discussed here, activity is taken as an example. Figure 6 represents the level of affect that activity has on mood. Also, Figure 7 represents the level of affect that activity has on a listener's music selection.

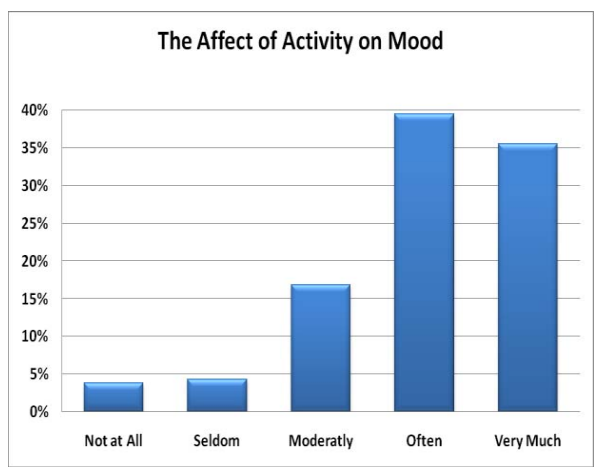

Figure 6: Represents the level of affect that activity has on a listener's mood

When one examines the general shape of both graphs, Figure 6 and Figure 7, it is noted that they are quite similar. For example in Figure 6, there is a significant step up from

Moderately to Often and then a gentle step down to Very Much. This movement is also repeated in Figure 7 with the exception of a scaling factor.

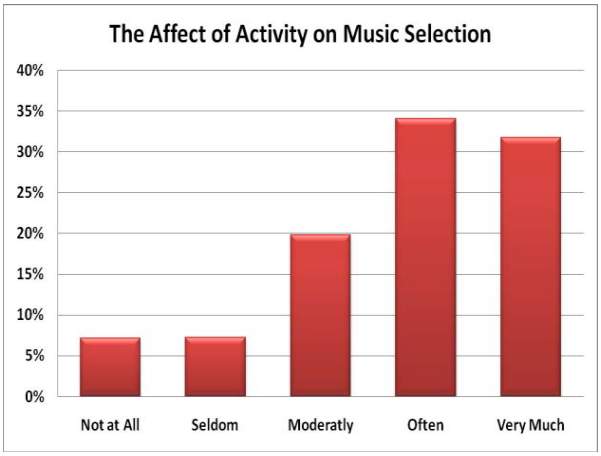

Figure 7: Represents the level of affect that activity has on a listener's music selection.

From the given example, it is suggested that instead of trying measure a listener's mood directly to select appropriate music for the listener, that simply measuring the listener's activity may yield the same outcome.

Further similarities between the listener's mood and their environment have also been noted from the survey results. The location of the listener being another. In this case, it is also suggested that instead of trying measure a listener's mood, that simply measuring the listener's activity may yield a similar outcome.

6. CONCLUSION

This paper has discussed the problem of access associated with large music collections. This problem is primarily due to the significant advancements in audio and storage technology. To counter act this problem, several techniques are available to allow a listener to automatically generate a music playlist from a large music collection. However, these processes are limited because the listener's mood is ultimately ignored.

It is suggested that environmental features, such as location, activity, temperature, lighting and weather have great potential as metadata and hence may be used to create a personalised automatic playlist generator for large music collections. Also suggested is that environmental metadata may be used to infer a listener's mood. For example, it was shown that a listener's activity seemed to affect a listener's mood and music selection equally. Similar conclusions can be drawn in regards to a listener's location.

To conclude, this paper has indicated that mood determines the listener's music selection process and demonstrated that an individual's environment strongly influences mood and hence the listener's selection process. Based upon these strong influences, it is concluded that environmental features pertaining to a listeners environment has significant potential as metadata and may provide a valuable resource in the automatic generation of music playlists.

7. ACKNOWLEDGEMENTS

This research was supported by Enterprise Ireland under the Commercialisation Fund.

8. REFERENCES

[1] Van Noorden L., "A Guide to European Funding", Keynote of *Audio Mostly: Conference on Interaction with Sound*, Fraunhofer Institute, Sept. 2007

[2] METRO, "MP3 Players Could Fit Millions of Songs", *METRO Newspaper*, AES, pp 9, 16th April 2008

[3] Gasser M., Pampalk E., and Tomitsh T., "A Content-Based User-Feedback Driven Playlist Generator and its Evaluation in a Real-World Scenario", Proceedings of *Audio Mostly: Conference on Interaction with Sound*, Fraunhofer Institute, pp 97-99, Sept. 2007

[4] Kravtsova N., Hollemans G., and Denteneer T.J.J., "Improvements in the Collaborative Filtering Algorithms for a Recommender System", Technical Note NL-TN-2001/542, Philips Research, Eindhoven, 2004

[5] AES, "Demystifying Audio Metadata", *Journal of the Audio Engineering Society Vol.32, No.2*, AES, 2003

[6] AES, "Intelligent Audio Environments", *Journal of the Audio Engineering Society Vol.55, No.10*, AES, pp 889-895, Oct. 2007

[7] Leue I., and Izmirli O., "Tempo Tracking with a Periodicity Comb Kernal", The 7th International Conference on Music Information Retrieval, Victoria, Canada, ISMIR 2006

[8] Greenwald AG., Brock TC., and Ostrum TM., "Psychological Foundations of Attitude", Academic Press, New York, 1973

[9] Covey SR., "The 7 Habits of Highly Effective People", *Simon&Schuster,* UK, 2004

[10] McDonald R., "Everyone is Musical", Keynote of *Audio Mostly: Conference on Interaction with Sound*, Fraunhofer Institute, Sept. 2007

[11] Toolos M., Tato R., and Kemp T., "Mood-Based Navigation through Large Music Collections of Musical Data", The 5th International Conference on Music Information Retrieval, Barcelona, Spain, ISMIR 2004

[12] Reynolds G., Barry D., Burke T., and Coyle E., "Towards a Personal Automatic Music Playlist Generation Algorithm: The Need for Contextual Information", Proceedings of *Audio Mostly: Conference on Interaction with Sound*, Fraunhofer Institute, pp 84-89, Sept. 2007