A Review of Factors Affecting Music Recommender Success

Alexandra Uitdenbogerd and Ron van Schyndel
Department of Computer Science, RMIT University
GPO Box 2476V, Melbourne 3001, Australia
+613 9925 2246
{alu,ronvs}@cs.rmit.edu.au

ABSTRACT

Much research has been published on musical taste, however, little has been studied by the builders of music recommenders. Implicit and explicit collaborative filtering has been used for making recommenders, in addition to the automatic classification of music into style categories based on extracted audio features. This paper surveys research into musical taste, reviews music recommender research, and outlines promising directions. In particular, we learned that demographic and personality factors have been shown to be factors influencing music preference. For mood, the main factors are tempo, tonality, distinctiveness of rhythm and pitch height.

1. INTRODUCTION

In the past twelve years, there has been interest in the development of techniques that provide personalised content to users. The type of applications have included filtering of news messages, presenting lists of stories or artwork that a user may be interested in, and so on. Most of these applications have applied a technique known as “collaborative filtering”. This involves collecting other users' opinions of how good or useful an item is, and then ranking items based on this information for presentation to the user.

While it may be argued that there has been some success with this technique, there is much room for improvement. Parallel to the development of collaborative filtering has been content-based filtering. This is an approach that tries to extract useful information from the items of the collection that are good indicators of their usefulness for a user. It is closely related to the field of information retrieval, which aims to develop better techniques to locate documents that satisfy a user's information need.

Currently, most music recommender services are based on editorial data, recommendations gleaned from the Internet user community, and browsing patterns. However, it is recognised that current approaches have important limitations, including inadequate raw data (in the case of editorial information), lack of quality control (in the case of user preferences), and lack of user preferences for new recordings. Deriving features from the music itself, rather than relying on customer behaviour is particularly important for introducing new music. A recommender system would never suggest new artists based only on customer behaviour, if no customer ever initially selected the new artist. Another limitation is the way the recommendations are presented: most systems use no more than a simple list of recordings. Further, there has been little effort to use knowledge from music psychology research to inform the choice of features to extract from audio, or to filter music selections.

In this paper, we rephrase the research question with a specific focus on music recommender systems, and consider the main factors that would affect the success of music recommender technology. We review work from a variety of sources, including research from the fields of psychology and marketing that relates to musical taste. After addressing these main factors and discussing the conclusions reached from the various branches of research related to this problem, we distill a set of guidelines as well as questions that remain to be resolved. We relate this to the aims of our research project.

2. THE PROBLEM

From the user's point of view, the purpose of a music recommender system is to recommend music that the user will be interested in. In order for the user to want to use the system it must be simple to use, with a minimum of input required from the user. Alternatively, there must be a clear and obvious incentive to the user that more effort in providing input will lead to better recommendations. The user may want to retrieve music based on preferences, style or mood.

3. PSYCHOLOGICAL FACTORS AND MUSICAL TASTE

When designing a system that recommends music, it is useful to learn from existing research into factors that affect musical taste. In this section we survey research that has occurred in the last eighty years on musical preferences. Most of the results come from work by social psychologists, but some comes from the more applied field of demographics for marketing.

Some of the research cited published before 1950 showed cultural biases. In addition there was a strong bias against popular music, for example, one author defined it as “music that is ranked by critics as tawdry, banal, and insipid” [9]. Researchers believed that one important purpose of musical education was to *improve* a student's musical taste. However, the experiments and their results appear to be sound and can be used as a starting point for experiments more targeted to music recommender systems.

3.1 Personality, Demographics and Music Preference

It has been shown that certain aspects of personality are correlated with music preference. In research published in 1939, Burt used the introvert versus extravert and stable versus unstable personality tests devised by Eysenck and concluded that stable extraverts prefer solid predictable music, stable introverts the more cognitive of classical and baroque styles, unstable extraverts the romantic styles expressing overt emotions, and unstable introverts the more mystical and impressionistic romantic works (discussed in [19] and [20]). More recently, it has been shown that the level of aggressiveness correlates with musical style, with more aggressive people being more likely to enjoy heavy metal or hard rock music [19].

Studies of different cultural groups, show different distributions of musical preferences. For example, Japanese adolescents have a higher likelihood of enjoying classical or jazz music than their American counterparts [33]. The study also concluded that there was generally a dislike of heavy metal due to the social connotations associated with the style.

Studies of age and demographics have shown that people prefer music that they were exposed to at a critical period of their life culminating at the age of about 23.5 [14].

One early study examined tempo preferences. These were shown for certain occupational groups to match the tempo of the occupation (discussed in Farnsworth [6] Chapter 1). For example, dressmakers preferred a moderately slow tempo, whereas typists preferred a fast tempo.

A study by Schuessler [29] showed that socio-economic, age and sex differences are correlated with music preference differences. In particular, his study of music preference focusing on people in Indiana showed that upper class women were more likely to enjoy classical music whereas working class men were more likely to enjoy hillbilly music. He was unable to rank the factors in order of importance as different factors were more influential for different pieces of music. Another observation that was made was that the more familiar a piece, the more likely it was to be enjoyed. While most of these observations probably still apply today, the musical genres may differ.

There have been studies that collected data on participants' attitude to music, ranging from general disinterest to spending large portions of the day listening to, or playing music [6]. An individual's attitude to music is likely to be related to their musical preference, however, we found few studies that explored this relationship. Schuessler's study [29] showed that women were more interested in music than men, as well as having somewhat different taste. However, it would be difficult to separate the various factors of gender, attitude, cultural issues, and personality from this study.

3.2 Perceived Music Quality

Farnsworth has shown that there is a consistency in how people rank classical music in terms of quality [6]. Further studies have shown that the same applies to popular music (discussed in [12]).

When it comes to deciding on the enjoyment of a piece of music, there are factors that have little to do with the sound of the piece itself. A study from around the time of World War II that associated the labels "romantic", "Nazi", or applied no label to the music, affected the listeners' enjoyment of the music. Similarly, when a revered composer's name was associated with a piece of music it was more enjoyed than when a less known composer's name was used (discussed by Farnsworth [6]).

3.3 Perception of Style and Mood

There are many different styles of music. The AllMusicGuide has 531 genres, Amazon 719, and MP3.com 430 [25]. When it comes to mood, however, the categories are not quite so numerous. Hevner devised 8 clusters consisting of 67 moods in total, which were then modified by Farnsworth based on experimental evidence into 10 clusters of a total of 52 moods, with one mood occurring in two clusters [6]. Perception of mood is fairly consistent within subcultures, but can vary for different cultural groups. Individuals may differ in their perception due to personal experience however.

The features that seemed to best distinguish mood for a set of pieces were tempo, tonality (for example major or minor key), distinctiveness of rhythm, and pitch (high or low). For example, solemn music tends to be slow, with a definite rhythm, and at a low pitch [6].

Style, however, has been classified according to a variety of parameters. Pachet and Cazaly found that classifications were based on genealogical, historical, geographical, functional, instrumental, as well as other enumerated types [25].

3.4 Other Human Factors

It has been shown that users are willing to provide more information in return for better quality recommendations [31], thus a short questionnaire that provides key personality indicators may be beneficial in fine-tuning recommendations. Indeed, there is at least one recommender web-site where there is a thriving community of users that regularly add ratings, discuss issues related to the site in a dedicated newsgroup, and play with the data available, for example, determining who their neighbours are by using a unique fingerprint [15]. Some users of this site have over a period of years entered more than one thousand ratings of stories.

3.5 Measurement of Relevant Factors

Most studies that assessed personality or preference relied on questionnaires. Many were of the multiple choice variety. Other tests are reversed, for example, it has been proposed to use musical preference to determine personality profile [6].

Personality and other psychology tests can easily be automated and delivered on-line, allowing extra evidence to be compiled for use in recommenders, for those users who wish to fine-tune recommendations.

3.6 Discussion

We can see from this survey of factors affecting music preference, that the following details about an individual may be useful for producing music recommendations:

- age
- origin
- occupation
- socio-economic background

- personality factors:
  - stability/instability (neuroticism)
  - introversion/extraversion
  - aggressive/passive
- gender
- musical education
- attitude toward music
- familiarity with the music or style

In the absence of large collections of data, these factors can be used to improve recommender precision.

Certain aspects that, to our knowledge, have not been addressed in psychological or other studies are the preferred level of complexity of music. The general consensus, supported by some studies, is that people will remain interested in music of greater complexity for longer. Thus, simple popular music may be appreciated with few listenings, but the listener tires of it sooner. We could extrapolate from an individual's preference for classical or popular music whether they prefer complex or simple music, but there are many levels of complexity to be found within each of these.

Another factor that doesn't appear to have been studied is the relative importance of the lyric content to the musical content. Through anecdotal evidence, some people will decide whether they enjoy a song largely on the lyric content, while others won't even notice the text, having all their attention focused on the music.

4. TECHNIQUES APPLICABLE TO RECOMMENDERS

Answering queries such as, "I like the song 'I've seen that face before'; what else is there that is similar?", "I want a piece of music for my film, that expresses defiance", "I like this kind of music, what else do you think I would like?" is the aim of designers of recommender systems.

The first query would probably need to be entered as an audio fragment of the original piece, and compared with a database for stylistic and mood similarities. The second would require a mood index, and the third would probably rely on the likes and dislikes of other users.

In this section we review the techniques of content, feature and collaborative filtering that are related to answering these types of musical query.

4.1 Collaborative Filtering

Collaborative filtering consists of making use of feedback from users to improve the quality of material presented to users. The feedback gathered can be explicit, in the form of user ratings and annotations, or implicit, such as the time users spend examining the content [5].

Early work in Collaborative Filtering published by Goldberg et al. [11] described the Tapestry system which filtered mail and news for users based on feedback from other users, but required the user to explicitly specify filters. Developers of the Grouplens project extended the technique by automating the filter based on similarity between a user's ratings and those of other users [24].

This idea has since been applied to a variety of types of information, including art-work [21], stories [15] and movies [13]. Recently Chen and Chen [2] have implemented a music recommender system that uses a form of collaborative filtering as well as features extracted from MIDI data to cluster pieces according to style in order to recommend music.

As collections of users and ratings get larger there may also be issues relating to efficiency as well as how effective a collaborative filtering approach is. The technique's effectiveness may also differ for different types of item. It is not clear whether a technique that works well for news or stories will be the best approach for music. It will be interesting to explore this question.

One problem with collaborative filtering is that new entrants in a music database do not get accessed by existing users, which means that they will not be selected for recommendation. While this problem can be addressed somewhat by having a "new music" section available to users, it will not influence the user who bypasses this section.

4.2 Content-Based Filtering

Much of the research in the field of audio analysis has been brought to bear in analysing music. A general and introductory treatment of the field was written by Roads [27]. The most significant problem in music analysis is to separate the musical semantics from the raw signal, which may have had many special effects (such as echoing) applied to it.

In feature-based filtering, a number of features are determined from the sampled music. These features involve such signal-processing methods as spectral and cepstral analysis and the use of instrument profiles. These parameters or feature combinations might be selected or determined automatically as a result of a training run, where the user's response to data with known attributes is used to model preferences.

Content-based filtering, a special case of feature-based filtering, is where we define a technique that relies on the determination of *semantic* attributes of the data. In this case, the feature elements are semantic components of the music. This is the more common approach to music analysis.

In this context, the word semantic can be applied in two ways: either as a progression of low-level musical symbols such as notes and rhythm elements; or at a higher, more general level where the features used are those attributes that are readily recognisable as the adjectives that people might use to describe and classify that piece or component of music - the 'defiance' of the question at the start of this section.

An ideal data-driven filter would be entirely content-based. That is, one where the features used by humans to classify and index music would be employed directly by the computer-based classification. The difficulty with constructing such a system lies in the requirement that the music be parsable by computer [30] (part of the motivation to use collaborative filtering was to have the user do the classification). In practice, most implementations use a combination of feature and content-based filtering [27].

The music information retrieval field of research has yielded systems that allow users to find a piece of music by humming a fragment of the tune [1, 10, 17, 23] (some of these systems are deployed on the web) but while most of this research has focused on melody matching, some other music features have been used to match pieces, such as rhythm [3], chords [4], and structure [7, 8]. For exact identification of recordings, the use of extracted audio features to create a unique fingerprint is also being explored.

Other music-related technologies have also been developed, such as techniques for automatic transcription of recorded music [18, 28] and signal processing for identification of instrumentation [18, 22]. A few techniques have been applied to recommender systems in the form of genre classification [32, 34].

In particular, Welsh et al. [34] described a system for performing similarity queries from a music database based only on characteristics of the music itself. The size of the feature space makes it impractical for users to provide the feature parameters directly, so users are required to provide a piece of music to query against. This provides an initial set of feature parameters and the musical piece may come from the database itself. Their algorithm essentially derives 1248 feature dimensions for each song, with feature parameters including a time-varying frequency histogram, the level of perceived "noise" present in the music, volume changes, and tempo and rhythm. These features are derived from 10-15 second samples at various stages in the music. A k-nearest neighbour matching approach was used to generate the best matches to a user's audio query. The approach was moderately successful in classifying music into one of seven genres, with classical music being the easiest to identify.

More recent work in genre classification has made use of musical surface features, such as the number of times per second the amplitude of an audio waveform changes between positive and negative (zero crossings), a measure of spectral shape (rolloff), as well as rhythm features revealed by wavelet transforms [32]. This was quite successful in separating several music genres. Again, classical music was easily identified, along with hiphop. Interestingly, the music surface features were slightly better indicators of genre than rhythm, but combining both feature sets worked best.

It is also possible to use the psycho-acoustic models of human perception developed for sound compression to analyse the music content. Since these models separate the perceived content from the rest, they can be used to reduce the data size for matching. This technique could allow music matching using MP3 format files to be performed without decompression. Jayant [16] has a general discussion about signal compression based on models of human perception, which makes it possible to discern patterns based on the *perceptual* rather than the *physical* properties of sound. Painter [26] offers a more exhaustive treatment of perceptual coding of digital audio. Both these papers and that of Roads [27] provide a large list of references.

5. RECOMMENDATIONS FOR RECOMMENDERS

From a research perspective there are many questions that can be explored in order to better understand techniques that will improve music recommenders. For collaborative filtering approaches we need to determine which of the currently tried methods of combining user ratings is optimal. Should data mining techniques be used, or k-nearest neighbour, or something else? Given knowledge about which techniques work, how can they be implemented efficiently?

Personal information from users, such as personality, age, origin and occupation can be combined with collaborative filtering to improve nearest neighbour estimation. For example, as it has been shown that music that one listened to at a specific age is more likely to be enjoyed, and that musical preference is socially conditioned, the music of other people of the same age may have more overlap than users of disparate ages.

For mood, the features that are most important are tempo, tonality, distinctiveness of rhythm and pitch height. Tempo can be extracted from audio using established techniques. Tonality as a specific feature hasn't directly been used, although the tonal histogram and tonal transition features extracted by Welsh et al. [34] may be well correlated to this. Distinctiveness of rhythm can be determined via the certainty associated with tempo extraction, that is, how clear the beats are. It doesn't appear to have been used as a feature yet. Melody pitch height is likely to be a difficult feature to extract from polyphonic music due to the confusion with harmonics associated with instrumental timbres, but by using peak values it may be possible to determine this.

Accurate genre classification via audio is more problematic than mood classification as there are so many genres in typical classification hierarchies. Based on existing categories, it seems that the main features that would aid in determining genre or style include rhythmic pattern, tempo, tonality, noise and instrumentation. Welsh et al. found that their noise feature was very closely associated with instrumentation of a piece of music.

Our research presently involves combining the above techniques and those of Welsh et al. [34] to allow recommenders to also classify and recommend new music.

6. SUMMARY

Through our survey of music psychology we have isolated factors that have been shown to affect human music preference. The main individual factors include aspects of personality, age, ethnicity, and socio-economic background. Preference is also influenced by what individuals believe about the music they hear, and their familiarity with it. Preference is influenced by the desire to be accepted by the group. Some of these features can be incorporated into a music recommender system via the use of questionnaires. For mood or style queries, some features that existing systems use, match those found via psychology research to be important indicators; however some other factors, such as distinctiveness of rhythm, do not appear to have been used.

7. ACKNOWLEDGEMENTS

We thank James Thom, Justin Zobel, Lin Padgham, and Santha Sumanasekara for their involvement in the early stages of this project. We also thank the anonymous referee for their helpful comments. This work has been partly supported by an RMIT School of Computer Science and Information Technology department grant.

8. REFERENCES

[1] J. Borchers and M. Muhlhauser. Design patterns for interactive musical systems. *IEEE Multimedia*, 5(3):36–46, 1998.

[2] H.-C. Chen and A. L. P. Chen. A music recommendation system based on music data grouping and user interests. In *Conference on Information and Knowledge Management*, pages 231–238, Atlanta, Georgia, USA, November 2001.

[3] J. C. C. Chen and A. L. P. Chen. Query by rhythm an approach for song retrieval in music databases. In *Proceedings of IEEE International Workshop on Research issues in Data Engineering*, pages 139–146. IEEE, 1998.

[4] T.-C. Chou, A. L. P. Chen, and C.-C. Liu. Music databases: Indexing techniques and implementation. In *Proceedings IEEE International Workshop in Multimedia DBMS*, 1996.

[5] M. Claypool, P. Le, M. Wased, and D. Brown. Implicit interest indicators. In *ACM International Conference on Intelligent User Interfaces*, pages 33–40, Santa Fe, NM USA, January 2001.

[6] P. R. Farnsworth. *The Social Psychology of Music.* Holt, Rinehart and Winston, New York, 1958.

[7] J. Foote. Visualizing music and audio using self-similarity. In *Proc. ACM International Multimedia Conference*, pages 77–80, Orlando Florida, USA, October 1999.

[8] J. Foote. ARTHUR: Retrieving orchestral music by long-term structure. In D. Byrd, J. S. Downie, T. Crawford, W. B. Croft, and C. Nevill-Manning, editors, *International Symposium on Music Information Retrieval*, volume 1, Plymouth, Massachusetts, October 2000.

[9] S. K. Gernet. *Musical discrimination at various age and grade levels.* The College Press, Washington, 1940.

[10] A. Ghias, J. Logan, D. Chamberlin, and B. Smith. Query by humming — musical information retrieval in an audio database. In *Proc. ACM International Multimedia Conference*, 1995.

[11] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. Using collaborative filtering to weave and information tapestry. *Communications of the ACM*, 35(12):61–70, December 1992.

[12] D. J. Hargreaves and A. C. North, editors. *The Social Psychology of Music*, volume 1. Oxford University Press, 1997.

[13] J. L. Herlocker, J. A. Konstan, and J. Riedl. Explaining collaborative filtering recommendations. In *Proceeding on the ACM 2000 Conference on Computer Supported Cooperative Work*, pages 241–250, Philadelphia, PA, December 2000.

[14] M. B. Holbrook and R. M. Schindler. Some exploratory findings on the development of musical tastes. *Journal of Consumer Research*, 16(1):119, June 1989.

[15] D. Howell. Alexandria digital literature. Web-site. Last accessed August 2002.

[16] N. Jayant, J. Johnston, and R. Safranek. Signal compression based on models of human perception. *Proc. IEEE*, 81(10):1385–1422, October 1993.

[17] T. Kageyama, K. Mochizuki, and Y. Takashima. Melody retrieval with humming. In *Proc. International Computer Music Conference*, 1993.

[18] K. Kashino and H. Tanaka. A sound source separation system with the ability of automatic tone modeling. In *Proc. International Computer Music Conference*, pages 248–255, 1993.

[19] A. Kemp. *The Musical Temperament.* Oxford University Press, Oxford, 1996.

[20] A. Kemp. *Individual Differences in Musical Behaviour*, chapter 2, pages 25–45. Volume 1 of Hargreaves and North [12], 1997.

[21] A. Kohrs and B. Merialdo. Using category-based collaborative filtering in the active webmuseum. In *IEEE International Conference on Multimedia and Expo*, pages 351–354, New York, July 2000.

[22] K.D. Martin and Y.E. Kim. Musical instrument identification: A pattern-recognition approach, 1998. Paper read at the 136th meeting of the Acoustical Society of America.

[23] R. J. McNab, L. A. Smith, I. H. Witten, C. L. Henderson, and S. J. Cunningham. Towards the digital music library: Tune retrieval from acoustic input. In *Proc. ACM Digital Libraries*, 1996.

[24] B. Miller, J. Riedl, and J. Konstan. Experiences with grouplens: Making usenet useful again. In *Proceedings of the Usenix Winter Technical Conference*, January 1997.

[25] F. Pachet and D. Cazaly. A taxonomy of musical genres. In *Content-Based Multimedia Information Access Conference (RIAO)*, Paris, April 2000.

[26] Ted Painter and Andreas Spanias. Perceptual coding of digital audio. *Proc. IEEE*, 88(4):451–513, 2000.

[27] C Roads. *The Computer Music Tutorial.* MIT Press, 1996.

[28] E. D. Scheirer. *Music Listening Systems.* PhD thesis, MIT, Massachusetts, June 2000.

[29] K. F. Schuessler. *Musical Taste and Socio-Economic Background.* PhD thesis, Indiana University, Bloomington, Indiana, 1980.

[30] U. Shardanand. Social information filtering for musical recommendation. Master's thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1994.

[31] K. Swearingen and R. Sinha. Beyond algorithms: An HCI perspective on recommender systems. In *ACM SIGIR Workshop on Recommender Systems*, Tampere, Finland, August 2001.

[32] G. Tzanetakis, G. Essl, and P. Cook. Automatic musical genre classification of audio signals. In J. S. Downie and D. Bainbridge, editors, *International Symposium on Music Information Retrieval*, volume 2, Bloomington, Indiana, USA, October 2001.

[33] A. Wells and H. Tokinoya. The genre preferences of western popular music by japanese adolescents. *Popular Music and Society*, 22(1):41, 1998.

[34] M. Welsh, N. Borisov, J. Hill, R. von Behren, and A. Woo. Querying large collections of music for similarity. Technical Report UCB/CSD00 -1096, U.C. Berkeley Computer Science Division, November 1999.