

**This is an electronic reprint of the original article.  
This reprint *may differ* from the original in pagination and typographic detail.**

**Author(s):** Wohlfahrt-Laymann, Jan; Heimbürger, Anneli

**Title:** Content Aware Playlist Generation with Multi-Dimensional Similarity Measure

**Year:** 2016

**Version:**

**Please cite the original version:**

Wohlfahrt-Laymann, J., & Heimbürger, A. (2016). Content Aware Playlist Generation with Multi-Dimensional Similarity Measure. In H. Jaakkola, B. Thalheim, Y. Kiyoki, & N. Yoshida (Eds.), Proceedings of the 26th International Conference on Information Modelling and Knowledge Bases - EJC 2016 (pp. 336-346). Tampere University of Technology. Tampereen teknillinen yliopisto, Porin laitos. Julkaisu, 18.

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

# Content Aware Playlist Generation with Multi-Dimensional Similarity Measure

Jan Wohlfahrt-Laymann<sup>a</sup> and Anneli Heimbürger<sup>b1</sup>

<sup>a</sup>*University of Twente, The Netherlands*

<sup>b</sup>*University of Jyväskylä, Finland*

**Abstract.** Music players and cloud solution for music recommendation and automatic playlist creation are becoming increasingly more popular, as they intent to overcome the issue of the difficulty for users to find fitting music, based on context, mood and impression. Much research on the topic has been conducted, which has recommended different approaches to overcome this problem. This paper suggests a system which uses a multi-dimensional vector space, based on the music's key elements, as well as the mood expressed through them and the song lyrics, which allows for difference and similarity finding to automatically generate a contextually meaningful playlist.

**Keywords.** Music Analysis, Music Information Retrieval, Multimedia Database, Automatic Playlist Generation, Context-Aware Recommendation System

## 1. Introduction

In recent years, with the increasing popularity of mobile devices and online music streaming services, systems providing music access have become more portable and available. As it has been frequently noted, more options often make the decision process more difficult; with large, and growing music databases it is often difficult for a user to find music according to their impression and mood. A user's preference for music depends on multitude of reasons, such as mood, impression, and context; such as in what company they are, time of day and their current activity. Users frequently make use of music in order to elevate a particular mood or emotion, as they provide means for the articulation of feelings [3]. In addition, familiarity with a particular musical piece has been shown to increase the likeliness, that a user will enjoy a piece, and thereby affects a user's current preference in music [16], [17]. For these reasons, a computational database approach to music analysis, mood and context recognition for music search and recommendation is promising, in that it would help users to find fitting music through the creation of fitting playlists.

Several difficulties are imposed on the system, that the system proposed in this paper tries to overcome through new and combined solutions. For example, it is often not easy for these systems to correctly identify the mood of a user as many contextual factors will affect mood and emotional response and thereby also a user's liking of an item.

Existing solutions trying to overcome these issues, usually consider music analysis of key elements, such as tonality, frequency, tempo, etc., or the analysis of lyrics. Other approaches, similar to the system proposed in this paper, consider the analysis of lyrics

and music key elements, in combination for mood analysis, classification and representation. The system proposed in this paper provides a solution for automatic playlist and query generation, based on mood represented in lyrics and music features through music information retrieval. Based on these values the system is able to perform distance measurement and similarity finding for automatic query and context based playlist creation.

## 2. Related Work

Multiple studies on the analysis of music tonality and lyrics, as well as the representation and analysis of mood in music have been conducted, through lyric analysis, as well as music information, which have been taken into account in the development of the system. Existing solutions allow the analysis of music elements and mood analysis [21] and query generation by tonality [7], to aid in solving the issue of the difficulty in expressing user impression to find the right music. The system proposed by Imai [7] analyzes key elements in music and visualizes them as colors, to express mood.

Subjectivity lexica allow for the identification of text to a sentiment; they consist of words subjectively analyzed based on sentiment. Wilson, Wiebe and Hoffmann [23] in the creation of their subjectivity clue lexicon have identified the difficulties in recognizing sentiment and full contextual sentiment analysis in text. The lexicon used by the researchers is made publicly available online.

Extending on the study of sentiment analysis, by using Jaynce Wiebe's subjectivity clue lexicon, Oudenne et al. [15] present several algorithms and compare their results; the researchers have shown three challenges in the sentimental analysis of song lyrics, which they identify:

1. A song might contain negative lyrics, but end on a positive note
2. A song might contain positive and negative lyrics, but the interpreted stanza identifies a particular subjectivity
3. Positive emotions may be expressed through negative things and vice versa

The results of the study show, that sentiment analysis of song lyrics is not easy, which resulted in a lower accuracy in comparison to other sentiment analysis tasks.

As noted by Frith, "Songs words are not about ideas ("content"), but about their expression" [3], lyrics provide means for the articulations of feelings. Lyrics are often experienced in non-verbal dimensions, giving a greater relevance to musical key elements, such as rhythmic features [14], which has been taken into account in the system implementation.

The system by Dang and Shirai [2] uses a machines learning approach, for a music search engine, using Naïve Bayes and support vector machine classifiers to analyze the expression of mood in song lyrics, by sorting to the mood clusters representing an exciting, joyful, sad, funny and aggressive mood respectively. The results show the difficulties mentioned of sentiment analysis of lyrics, previously identified [15], and therefore render the system too unreliable for production systems.

For music analysis of key elements, Thayer's two-dimensional model of mood [20] has been used in a system [13], which shows mood responses from pleasant to unpleasant and quiet to energetic, also referred to as valence and arousal dimensions, sometimes also including a dominance or tension dimension. The music features intensity, timbre and rhythm are extracted, through audio analysis from the music file,

classified and expressed in the music mood clusters Contentment, Depression, Exuberance, and Anxious/Frantic.

The 5D World Map System [9] is a spatial-temporal semantic space of multimedia objects. The researchers realized a five dimensional multimedia map. The system consists of one temporal, one semantic and three spatial dimensions, which allows for cross-cultural and environmental understanding, through the analysis and visualization of environmental change. This allows for fast recognition of localized events and problems through tagging of images.

### 3. Mood Categorization

One of the first systems and most well-known taxonomy of mood in music is the circle of eight adjectives, created by Hevner [4]. The system has often been adapted and prominently used for the classification of music in studies on the field of mood and music. Hevner used adjectives, such as spiritual, melancholy, sentimental, serene, playful, cheerful, dramatic and empathic, along similar adjectives, to describe the six keyword clusters that can be used for music classification.

In comparison, Russell's model consists of 28 adjectives scaled in a circle with the dimension pleasure-displeasure and degree of arousal [18]. As concluded in the study by Hu [5], who compared the usage of the Hevner and Russell taxonomy with last.fm tags, these models often use outdated vocabulary and context-dependent finer distinctions are made; for this reason, classifiers should be context-, user-, and usage dependently adapted. The Thayer mood model for lyric based mood detection shows some similarities to Russell's model. The model by Thayer defines dimensions for stress and energy levels and defines four categories: anxious, depression, exuberance and contentment for categorization [20].

Kiyoki and Chen [8] made use of the Hevner mood classifiers in their system, for decorative multimedia creation. The researcher made use of the method earlier described by Kiyoki et al. [10] for the creation of impression metadata to music data. Mood was visualized through the usage of color, which allowed for a time-dependent mood analysis in music.

As previously shown, music impression or mood classification systems, show a variety of approaches for taxonomies and mood analysis approaches that are context dependently adapted and used. However, context-aware recommendation systems should consider more context dependent variables as they affect the listener's mood and thereby enjoyment of a song. A music context recommendation system has been described by Baltrunas et al. [1], who used five dimensions as context in their context-aware recommender systems: activity, weather, time of day, 'valence' mood, and 'arousal' mood. 'Valence' mood ('happy', 'sad'), and 'arousal' mood ('calm', 'energetic') can be loosely compared to the 'valence' and 'arousal' dimensions of Russell's model. A similar system [25] demonstrated the effect of day, location and companion as contextual variables in their context-aware recommendation system for movies.

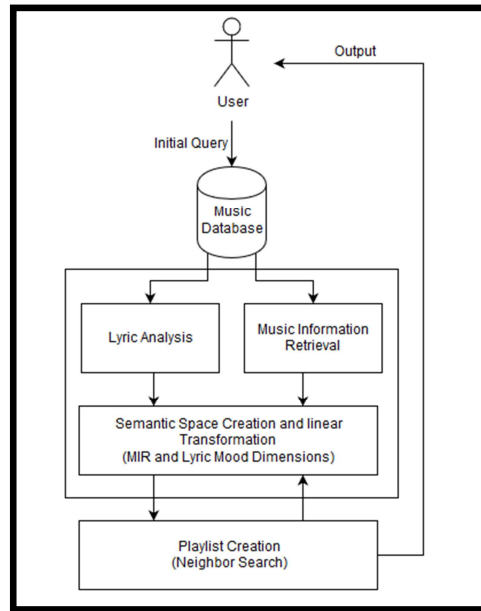
One of the difficulties in real world scenarios is acquiring more contextual information to be used in context-aware recommendation systems for more accurate predictions, many systems acquire this information, through user questionnaires, which often has a high risk of biases from contextual information that is not measured. Another more intrusive system [12] makes use of measurements from the user's

heartbeat in their playlist recommendation system. With the increasing availability, portability and access of biosensors in the Internet of Things, future recommendation system can easily acquire and use more contextual information in their recommendations.

## 4. System Implementation

### 4.1. Basic Operation

The system allows a user to perform a query for a song the system will then return a playlist based on the mood perceived through lyrics and the music features retrieved through music information retrieval. Lyric and Music analysis is performed on the music database in order to create a hyperspace. The system will perform a neighbor search starting with the search query and then continuing to move through a hyperspace of the retrieved data. Figure 1 shows a diagram of the basic system setup.



**Figure 1.** Software Diagram

### 4.2. Music Feature Analysis

Music mood, impression or emotion analysis has been performed through the use of various audio features. Music features are commonly described in four categories: intensity, pitch, rhythm and timbre. Imai et al. [7] analyzed tonality, by applying the Krumhansl-Schmuckler key-finding algorithm, in their mood analysis and visualization of music files. The system allows a user to input a music metadata query to receive search results visualized by tonality.

Similarly another system by Trang et al. [21] uses tonality with a culture-dependent transformation matrix to generate impression metadata for different music cultures. The resulting system therefore allows for music retrieval and interpretation for different cultures, based on their impression interpretation of music tonality.

In their personalized music filtering system, Kuo and Shan [11] extract melody and perform analysis of user preference in their recommendation system.

While these systems use MIDI-files, which allow for retrieval of information about the music's features, such as tonality, at relative ease, other approaches perform analysis on audio files, such as MP3. Audio Analysis, such as spectral analysis, Tempo, Transcription, Tonality, and Structure can be performed at relative ease, through the use of frameworks and platforms, such as jAudio<sup>1</sup>, MARSYAS<sup>2</sup> and the MIR toolbox<sup>3</sup>.

The system presented in this paper makes use of the MIRToolbox for music information retrieval. The MIRToolbox includes a set of Matlab functions for the analysis of audio files. Information about the music features dynamics or intensity, pitch, rhythm, tonality and timbre is retrieved and stored in a comma-separated file. For the Intensity, the Root-Means-Square Energy is calculated through the use of the mirrms function, the function calculates the RMS value of the amplitude, therefore the global energy of the signal. The pitch is calculated with the mirpitch function, the pitch for the entire music file is computed, with an autocorrelation function of the waveform, the best pitch is selected and returned. The mirpitch function computes the tempo in beats per minute (BPM), frame by frame for the audio waveform and returns only the best beat. The best key for the audio track is computed from the mirkey function.

#### 4.3. Lyric Sentiment Analysis

Through the usage of natural language and subjectivity clue lexica previous studies have shown the possibility of realizing mood classification from specific text sources. However, the results of Dang and Shirai [2] and Oudenne et al. [15] have shown several difficulties in realizing effective and reliable classification. The issues in the classification of lyrics, identified by researchers, have been described, that the emotional response can often only be understood when an entire stanza is considered, and may express different emotions, when only one word or line is interpreted. These issues have been noted as the cause for the often unreliable usage of key words to express different emotions.

The combination of lyrics and music for mood classification has been shown, in one of the first studies on the field by Yang and Lee [24], who have tested for psychological features driving emotions in song lyrics. The researchers created 182 psychological feature vectors of the General Inquirer, in order to disambiguate emotion, due to an excessive vocabulary size for songs. By fusing acoustic and text features, the classification accuracy could be increased.

Lyric and other text sentiment analysis systems often make use of machine learning techniques, for example through the use of Naïve Bayes Classifier. Common approaches for the classification of lyrics use bags-of-words and n-gram representations. The use of content word n-gram features [6] has shown better results

---

<sup>1</sup> <http://jaudio.sourceforge.net/>

<sup>2</sup> <http://marsyas.info/>

<sup>3</sup> <https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>

for mood analysis and classification, than classification through audio features retrieved by MARSYAS. Furthermore, the researchers classified lyrics to 18 mood categories, of which seven lyrics features outperformed audio on classification. These results show the relevance of lyrics, when classifying lyrics.

Audio analysis and lyric features have previously been successfully used together for mood classification of music. Another system [19] makes use of both for classification with the help of support vector machines. The researchers classify music into different mood classes. Before processing the data the research make use of the Porter Stemmer algorithm after removing punctuation. In the computation of the average mood for an entire song, the classification results have shown that songs consisting of sections with opposing emotional features can average out, and be characterized with value 0 for that dimension, which is not the same as a song with no such emotional features. This is an issue that is certainly also plays a role in the system proposed in the paper.

Word list play an important role in sentiment analysis [26]. They are created through an opinion mining process and have found use cases in analysis, and especially machine learning. Word lists define words along one or more sentiment dimension, and are therefore very useful in the sentiment, mood or emotion analysis and classification of text.

For the lyric analysis, the sentiment analysis file, as described by Warriner, Kuperman and Brysbaert [22] was used. The file categorizes 13,915 English lemmas with a sentiment rating in the dimensions: valence, dominance and arousal. In accordance with the dimensions defined in Thayer's model for mood classification [20]. The data has been collected with Amazon Mechanical Turk<sup>4</sup>. The implementation of the lyric analysis and the final playlist generation system is written in C++. Before the analysis of lyric, the system applies stemming with the Porter Stemmer Algorithm to the lyrics and the word list, from the sentiment analysis results. This means the system will consider the stem of the word, instead of the words in their conjugated and plural form, as they might appear in lyrics. The valence, dominance and arousal values are computed from the word list with the mean values from the entire identifying group from Warriner et al. [22], by comparing the lyrics word by word and computing the average values for the text from all identified words.

#### *4.4. Final Analysis and Playlist Generation*

After retrieving lyrics and computing the audio and lyric analysis, for all music files, the data for RMS, Pitch, Tempo, Key and Inharmonicity from the audio and Valence, Arousal and Dominance from the lyric analysis are defined as dimensions for further processing. Linear transformation is performed on each dimension, so that the data is defined in an orthogonal hyperspace with a size of 0 – 10 in every dimension.

When the user performs a search query the system is able to automatically create a playlist beginning with the query results along the specified dimensions. Within the 8 dimensional space similar items are located close to each other. The system therefore creates a similarity search by creating an 8-dimensional sphere with its center as the location of the first music track. The values of the music tracks are defined as floating point the initial radius and increment of the sphere is therefore the smallest exponent of the data. Because the data is defined to the seventh exponent at most, the value of the

---

<sup>4</sup> <https://www.mturk.com/mturk/welcome>

initial radius is  $10^{-7}$ . The algorithm to query for the next track searches tracks for which the distance to the center of the sphere is smaller than the radius for every dimension. The operation will immediately break out when the distance is bigger or the file is already in the playlist, thereby speeding up the nearest neighbor search. At each iteration of the algorithm, where no results have been returned the radius of the n-sphere is incremented by the initial radius.

The system will create a playlist with 11 tracks in total, starting with the query track. The playlist is saved in the pls file format.

In order not to have too quickly changing playlist and making the playlist better fit to the current mood, the mood of the initial tracks. The following equation have been tested to generate the new center location of the sphere:

$$new_n = old_n + (0.4 \times 0.9^S \times distance)$$

$$new_n = old_n + (0.5 - 0.4 \times e^{-1.75 \times S}) \times distance$$

With  $S$  representing the current size of the playlist, the intention of the different equation is to represent different levels of change, depending on the size of the playlist. The first equation performs a slower change, by moving the center less in the beginning, thereby including more files in the area of the initial query track, and performing a faster change at the end. The second equation slowly increases the level of change to the limit of 0.5 for the movement of the center. However, for the chosen dataset both equation delivered good results with slight changes at the end of the playlist, usually on the order of the tracks. Bigger changes should be considered, when a bigger dataset is considered. For the test results of the system only the second equation is considered, because it retrieved results slightly faster.

Table 1 showcases a small extract of the code, the first function is used for the creation of the next center point from which the distances to the music tracks will be measured. The second function is called with the initial sphere center, which is the location of the queried track, as well as the number of playlist items to create, excluding the track, that has already been retrieved.

**Table 1.** C++ Code extract

---

```
vector<float> createNextSphere(vector<float> old_v, MusicTrack next) {
    vector<float> new_v;
    for (int i = 0; i < old_v.size(); ++i) {
        //new_v.push_back(old_v[i]+((0.4*(pow(0.9,Playlist.size())))*((next.DataResults[i]-old_v[i]))));
        new_v.push_back(old_v[i]+((next.DataResults[i]-old_v[i])*(0.5-(0.4*(exp(Playlist.size()*(-1.75)))))));
    }
    return new_v;
}

void CreatePlaylist(vector<float> Sphere, int i2c) {
    if (i2c != 0) {
        Logger(LogFile).logMessage({"Hypersphere center created"});
        float radius = 0;
        float increase = pow(10,-5);
        MusicTrack *Neighbor = nullptr;
        while (!Neighbor) {
```

---



---

```

radius += increase;
Neighbor = returnClosest(Sphere, radius);
if (radius > 10) {Logger(LogFile).logMessage({"ERROR"}); return;}
}
Playlist.push_back(*Neighbor);
Logger(LogFile).logMessage({"The closest Neighbor is:",Neighbor->title,"-",Neighbor->artist});
Sphere = createNextSphere(Sphere, *Neighbor);
CreatePlaylist(Sphere, i2c-1);
} else Logger(LogFile).logMessage({"Playlist Complete"});
}

```

---

## 5. Dataset

The dataset consists of 89 popular music tracks in the English language from 1965 – 2015. The music combines a variety of moods and encompasses a wide range of genres, including Electronic, Reggae, Metal, Hip-Hop, Rock, Country and more. The music all has fully English lyrics that are easily accessible online. Lyrics are retrieved from lyrics.wikia.com, metrolyrics.com and azlyrics.com which have a high accuracy for the tested dataset. After removing punctuation, the Porter Stemmer Algorithms is applied and the stemmed results are tested against the stemmed word list, to retrieve the average valence, arousal and dominance values. The music files are encoded in MPEG-1 Audio Layer 3 with bitrates in the range of 128 – 320 kBit/s and a sampling frequency in the range of 44100 – 48000 Hz. In order to simulate a realistic real-world music database, and avoiding tracks with a sampling rate too low to retrieve meaningful results.

## 6. Results

The implemented system is able to create mood and music context dependent playlist in pls file format with songs from the testing database, based on the identified characteristics from an initial track in the database. The system allows the creation of playlist relatively quickly, with files stored in common formats allowing for potential use cases of the system in real-world scenarios. Difficulties for the system are short lyric with too little variation and lyrics with a large number of metaphorical and analogical word, that overall represent a different impression or mood, when the entire stanza or song is considered as context.

The playlist results when querying “Temporary Home” by Carrie Underwood and Michael Jackson’s “Billie Jean” can be found in Table 2.

**Table 2.** Playlist results for different query tracks

Carrie Underwood - Temporary Home	Michael Jackson - Billie Jean
3 Doors Down - Here without you	Elton John - Can You Feel the Love Tonight
Katy Perry - Teenage Dream	Eric Clapton - Knockin' on Heaven's Door
Coldplay - Fix You	Breaking Benjamin - The Diary of Jane

Adele - Turning Tables	Bob Marley & The Wailers - Is This Love
Chet Faker - Gold	30 Seconds To Mars - This is War
Adele - Set Fire to the Rain	Beyoncé - Sweet dreams
Flight Facilities - Clair de Lune	30 Seconds To Mars - Hurricane
Rihanna - Russian Roulette	The xx - Crystalised
Christina Perri - Jar of Hearts	Damian Marley - Welcome to Jamrock
Miranda Lambert - The House That Built Me	Norah Jones - Don't Know Why

---

The results show the capabilities of the system to create mood and context dependent playlist. Problems in the identification and playlist creation are partially due to the previously identified difficulties in the analysis process of lyrics and music features, as well as the small testing database.

## 7. Conclusion

The results have shown the effective realization of a system capable of automatically creating mood based playlist on the basis of an initially queried music track. Systems for the automatic creation of playlist are very relevant today, as systems and services for their realization are becoming more accessible. Music preferences depend on a variety of factors, such as context, impression and mood, which requires a deeper recognition and understanding of contextual information from context-dependent recommender systems. This paper proposed a new way of playlist generation through context dependent analysis of music. The system was able to generate mood-based playlists with relatively good results on the tested dataset. Problems for the system, that lead to unreliable classification in lyrics were primarily the use of metaphors, similes, analogies and homographs, with a different meaning when the words are considered in their context, making it difficult for non-human interpreters, as well as lyrics with words of opposing meanings, to express a certain sentiment, that the system averages out.

## 8. Further Research

Further research should include direct querying methods for mood and translation methods from the music element analysis to mood values. While the primary requires a relatively easy step, for example by placing key words in a similar hyperspace for comparison, and the realization of the first query, or simply by adding mood keywords or tags from sites such as last.fm or allmusic.com for performing the query, as it has been realized in multiple studies, the latter will prove more difficult, if values should be kept as floating points within a hyperspace, rather than classifying music to mood categories by its elements, and reducing the resolution of the analysis results.

The chosen database has been relatively small, when issuing the system for a large database, better resources for lyric acquisition should be found, which has proven relatively difficult, due to copyright reasons; in addition, the implementation of algorithms such as k-d tree or even locality sensitive hashing should be considered for faster nearest neighbor search. Hashing could be implemented to speed up the search query operation. However, due to the small dataset and algorithmic efficiency, playlist results could be retrieved relatively quickly.

## References

- [1] Baltrunas, L., Kaminskas, M., Ricci, F., Rokach, L., Shapira, B., & Luke, K.-H. (2010). Best usage context prediction for music tracks. In *Proceedings of the 2nd Workshop on Context Aware Recommender Systems*. Retrieved from [http://www.inf.unibz.it/~ricci/papers/CARS2010\\_best\\_context.pdf](http://www.inf.unibz.it/~ricci/papers/CARS2010_best_context.pdf)
- [2] Dang, T.-T., & Shirai, K. (2009). Machine Learning Approaches for Mood Classification of Songs toward Music Search Engine. In *International Conference on Knowledge and Systems Engineering, 2009. KSE '09* (pp. 144–149). <http://doi.org/10.1109/KSE.2009.10>
- [3] Frith, S. (1998). *Performing Rites: On the Value of Popular Music*. Harvard University Press.
- [4] Hevner, K. (1936). Experimental Studies of the Elements of Expression in Music. *The American Journal of Psychology*, 48(2), 246–268. <http://doi.org/10.2307/1415746>
- [5] Hu, X. (2010). Music and mood: Where theory and reality meet. Retrieved from <http://www.ideals.illinois.edu/handle/2142/14956>
- [6] Hu, X., & Downie, J. S. (2010). When Lyrics Outperform Audio for Music Mood Classification: A Feature Analysis. In *ISMIR* (pp. 619–624). Retrieved from <http://ismir2010.ismir.net/proceedings/ismir2010-106.pdf>
- [7] Imai, S., Kurabayashi, S., & Kiyoki, Y. (n.d.). A Music Database System with Content Analysis and Visualization Mechanisms.
- [8] Kiyoki, Y., & Chen, X. (2009). A semantic associative computation method for automatic decorative-multimedia creation with Kansei information. In *Proceedings of the Sixth Asia-Pacific Conference on Conceptual Modeling-Volume 96* (pp. 7–16). Australian Computer Society, Inc. Retrieved from <http://dl.acm.org/citation.cfm?id=1862743>
- [9] Kiyoki, Y., & Chen, X. (2014). Contextual and Differential Computing for the Multi-Dimensional World Map with Context-Specific Spatial-Temporal and Semantic Axes. *Information Modelling and Knowledge Bases XXV*, 260, 82.
- [10] Kiyoki, Y., Wangler, B., & Jaakkola, H. (2005). *Information Modelling and Knowledge Bases XVI*. IOS Press.
- [11] Kuo, F.-F., & Shan, M.-K. (2002). A personalized music filtering system based on melody style classification. In *2002 IEEE International Conference on Data Mining, 2002. ICDM 2003. Proceedings* (pp. 649–652). <http://doi.org/10.1109/ICDM.2002.1184020>
- [12] Liu, H., Hu, J., & Rauterberg, M. (2009). Music Playlist Recommendation Based on User Heartbeat and Music Preference. In *International Conference on Computer Technology and Development, 2009. ICCTD '09* (Vol. 1, pp. 545–549). <http://doi.org/10.1109/ICCTD.2009.246>
- [13] Lu, L., Liu, D., & Zhang, H.-J. (2006). Automatic mood detection and tracking of music audio signals. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(1), 5–18. <http://doi.org/10.1109/TSA.2005.860344>
- [14] Moser, S. (2007). Media modes of poetic reception. *Poetics*, 35(4-5), 277–300. <http://doi.org/10.1016/j.poetic.2007.01.002>
- [15] Oudenne, A. M., Swarthmore, P. A., & Chasins, S. E. (n.d.). Identifying the Emotional Polarity of Song Lyrics through Natural Language Processing. Retrieved from [http://www.sccs.swarthmore.edu/users/11/aoudenn1/Ashley\\_M.\\_Oudenne/Research\\_files/emotionalPolarity.pdf](http://www.sccs.swarthmore.edu/users/11/aoudenn1/Ashley_M._Oudenne/Research_files/emotionalPolarity.pdf)
- [16] Peretz, I., Gaudreau, D., & Bonnel, A.-M. (1998). Exposure effects on music preference and recognition. *Memory & Cognition*, 26(5), 884–902. <http://doi.org/10.3758/BF03201171>
- [17] Rentfrow, P. J., Goldberg, L. R., & Levitin, D. J. (2011). The Structure of Musical Preferences: A Five-Factor Model. *Journal of Personality and Social Psychology*, 100(6), 1139–1157. <http://doi.org/10.1037/a0022406>
- [18] Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178.
- [19] Schuller, B., Dorfner, J., & Rigoll, G. (2010). Determination of Nonprototypical Valence and Arousal in Popular Music: Features and Performances. *EURASIP Journal on Audio, Speech, and Music Processing, 2010*, 1–19. <http://doi.org/10.1155/2010/735854>
- [20] Thayer, R. E. (1989). *The Biopsychology of Mood and Arousal*. Oxford University Press.
- [21] Trang, N. N., Sasaki, S., & Kiyoki, Y. (n.d.). A cross-cultural music museum system with impression-based analyzing functions.
- [22] Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45(4), 1191–1207. <http://doi.org/10.3758/s13428-012-0314-x>

- [23] Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing Contextual Polarity in Phrase-level Sentiment Analysis. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing* (pp. 347–354). Stroudsburg, PA, USA: Association for Computational Linguistics. <http://doi.org/10.3115/1220575.1220619>
- [24] Yang, D., & Lee, W.-S. (2004). Disambiguating Music Emotion Using Software Agents. In *ISMIR* (Vol. 4, pp. 218–223). Retrieved from [https://courses.ischool.utexas.edu/Winget\\_Megan/MeganCollection2/files/10952/Yang\\_MusicEmotionSoftwareAgents.pdf](https://courses.ischool.utexas.edu/Winget_Megan/MeganCollection2/files/10952/Yang_MusicEmotionSoftwareAgents.pdf)
- [25] Zheng, Y., Mobasher, B., & Burke, R. D. (2013). The Role of Emotions in Context-aware Recommendation. *Decisions@ RecSys, 2013*, 21–28.
- [26] Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *FNT In Information Retrieval*, 2(1–2), 1-135. <http://dx.doi.org/10.1561/15000000011>