

A Review on Song Recommendation Techniques

Ms. Samruddhi S. Doshi¹, Dinesh B. Hanchate²

¹M.E. Second Year Student, Department of Computer Engineering, VPKBIET, Baramati, Maharashtra, India

²Department of Computer Engineering, VPKBIET, Baramati, Maharashtra, India

Abstract - Singing skills of each singer help in choosing accurate songs for them. These skills are determined by using vocal competence. Matching a song to singer's vocal competence is the major and difficult task in song recommendation. There are other song recommendation systems, which works on recommending songs of listener's interest. These traditional approaches did not consider the singer's ability of singing which leads to poor singing performance by singer. This paper focuses on the techniques related to music retrieval such as content based and collaborative. This paper has focused on the other method of song recommendation and the techniques of feature extraction, feature selection. Music transcription and song ranking using learning-to-rank method has explained.

Key Words: Singing Skills, Content Based Music Retrieval, Integrating Multiple Acoustic Features, Automatic Music Transcription, Feature Selection, Tree Structure Locality Sensitive Hashing, Social Singing Community.

1. INTRODUCTION

Singers can sing well as per their singing skills and the choice of accurate song is depend on the skill. For this purpose, vocal competence is used. Different singers are having different singing skills. Those singers give a better performance, only if, chosen song is suitable to singing skills of singers. At many a times, singers are depressed due to their own performance. The reason behind it is the bad choice of song, rather than the singing ability. e. g. it is difficult for a girl with soft voice to sing a song which require strong voice to express strong emotions[1]. To achieve good performance, songs are chosen such that, those songs were matching with singer's skills.

Traditional approaches of song recommendation research techniques in the field of content based song recommendation and uses collaborative techniques to recommend song. These techniques discovers favourite music of user in terms of music content similarity such as moods and rhythms. Collaborative techniques recommend a song in a group of users having the same interest. Content based music retrieval methods are given below. This paper discussed content based retrieval, ability based retrieval, evaluation of singer's performance, feature extraction, feature selection and ranking methods. These methods are useful and helpful for the understanding the need of making song recommendation according to ability of singer. Also the evaluation of singer's performance, feature extraction,

feature selection and ranking methods are useful in easily understanding procedure of recognizing singer's vocal ability.

This paper focuses on three methods of content based techniques. The first technique of Content Based Music Retrieval (CBMR) is given for retrieving songs according to singing preferences of user. This allows users to find new songs which they expected to like. Another technique of CBMR gives the system to construct music descriptors. That system supports efficient content-based music retrieval and classification. The next method of content based retrieval is, Scalable Content-Based Music Retrieval which depict Content-Based Music Retrieval Using Chord Progression Histogram and Tree-Structure LSH. The techniques of content based method only focused on the song recommendation as par the listener's interest. Consideration of the singer's ability is main aspect in song recommendation which has not covered in earlier studies. So for making song recommendation in social singing community, singing-song recommendation framework is used. Rather than recommending songs which people may like to listen, system used to recommend suitable songs that people can sing well. System is developed with the aim of finding the difficulty orderings of songs, from the song performance ratings of every user.

To evaluate the singer's performance, an automatic singing evaluation system for Karaoke performances has been used. The performance of singer evaluated in terms of technical accuracy and system then assign a rating score. The main task in music retrieval is feature extraction. One of the approach of feature extraction is used for the purpose of key detection. Basically, key in the music is the set of music notes which are primarily used for constructing a piece of music. The next method is about Automatic Music Transcription (AMT). Which is the process of conversion of an audio recording to some form of musical notation. The core problem for considered in AMT system was the detection of multiple concurrent pitches. This paper analyses limitations of current AMT methods and identified directions for future research. For selection of feature which helps in evaluating singer's performance, techniques are required. Next technique studied is feature selection method. This method has given a feature selection (FS) algorithm for filtering the low efficiency features for the fast speech emotion recognition. Next method studied is of the ranking which is used for retrieving songs matching with singer's vocal competence from database.

2. METHODS

2.1 Content Based Music Retrieval Methods:

In [2], K. Hoashi, K. Matsumoto, and N. Inoue gives Content Based Music Retrieval (CBMR) method. This method is used to retrieve songs as per the users singing preferences. This allows users to find out new songs which they expected to like, also given relevance feedback methods, which improves the performance of music retrieval method. The burden on users of inputting learning data to the system has been reduced with use of method which generate user profiles based on the genre preferences. Also they do refinement of such profiles based on relevance feedback. This music information retrieval method is based on TreeQ. Firstly, tree generation done on learning data samples, which consists songs from both the set of songs which a user likes (good songs), and which a user didn't like (bad songs). Vector representation done based on user preferences and this task is done by inputting the good songs from the learning data set through the tree, and calculating a histogram.

- **Advantages:**

- 1) System has considered user's musical preferences which are highly ambiguous.
- 2) Relevance feedback method is used to improve user's preferences.

- **Limitations:**

- 1) Worked on limited dataset.
- 2) System considers only listener's interest not the singer's vocal ability to retrieve specific song for recommending to singer. Due to this, singers do not give good performance in their singing.

Another technique of content based music retrieval is given by J. Shen, J. Shepherd, and Anne H. H. Ngu[3]. This method have used InMAF (Integrating Multiple Acoustic Features) framework which is fast as well as robust descriptor generation method. This approach combines multiple vocal feature vectors and musical perception given by the singer then it produces small single feature vector which enhance the music data retrieval and classification process. The work of generating music descriptor carried on two stages:

- a. Dimension reduction via Principal Component Analysis(PCA),

- b. Non-linear Neural Network (NN) used for generating music descriptors.

In step one, it extracts features such as Timber, Rhythm and Pitch having dimension vectors as 33, 18, 18 respectively. PCA is generally dimension transformation method which in second step used as pre-processing step for Neural Network. PCA provides features to neural network where dimension reduction of features has been done.

- **Advantages**

- 1) It captures higher level semantic concepts and present those features as low dimensional feature vectors.
- 2) System integrates newer lower-level acoustic features without any difficulty.
- 3) It improves efficiency of content-based music retrieval.
- 4) It gives effectiveness and robustness against various kinds of audio alteration.

The next method of content based music retrieval is Scalable Content-Based Music Retrieval[4]- This method is given by Y. Yu, R. Zimmermann, Y. Wang, and V. Oria who described Content-Based Music Retrieval Using Chord Progression Histogram and Tree-Structure LSH. As multimedia content increasing day by day over the Internet, music information retrieval is becoming a difficult. This method used the melody similarity for quick and reliable retrieval of relevant songs. The system works in two phases- firstly, music semantics used for compact and accurate representation of audio tracks. Then chord progressions has been recognized from audio signals based on trained music rules, and for improving recognition accuracy method has used multi-probing. Afterwards as a mid-level feature, concise Chord Progression Histogram (CPH) has computed from each audio track, which used to describe audio content. As a second step it did efficient organization of audio tracks according to their CPHs by using only one locality sensitive hash table with a tree-structure.

- **Advantages:**

- 1) Scalability compared with other methods.
- 2) Compared to other methods, the used algorithm has improved the accuracy of summarization and indexing. This algorithm have taken a further step for the optimal performance which is determined by an exhaustive sequence comparison.

- **Limitations:**

- 1) It explored techniques only in the domain of content based song recommendation. These techniques have only the purpose of discovery of user's favourite music. But, this method is not efficient because, the low-level features cannot fully represent the user's interests.

2.2 Methods considering Singer's Ability:

The techniques of content based method only focused on the song recommendation as per the listener's interest. Instead of recommending songs as per the listener's interest, there is need of recommending songs which are suitable to voice of particular singer. Consideration of the singer's ability is main aspect in song recommendation which is not covered in earlier studies. K. Mao, J. Fan, L. Shou, G. Chen and M. Kankanhalli [5] have proposed a song recommendation framework which could be used for making song recommendation in social-singing community. Despite of making song recommendation which people may like to listen, this system used to recommend suitable songs which are matching with singer's voice. So, in social singing community, when any singer uploads it's own recording, there are ratings based on singing performance of the singer. This helps in process of finding the song difficulty orderings from the song performance ratings of every user. Then system transforms the difficulty orderings into a difficulty graph. Iterative inference algorithm is used to recommend songs according to the difficulty graph.

- **Advantages:**

- 1) The difficulty orderings found in one Social Singing Community (SSC) can be reused and updated in another SSC system.
- 2) For music experts after giving sufficient resources, it is possible to build up an ever-growing song difficulty ordering database and then provide singing song recommendation services for any social singing community.
- 3) It takes into consideration somewhat singer's capability.

- **Limitations:**

- 1) It does not exact find the vocal ability of particular singer because it only considers the listener's review.

2.3 Method to Evaluate Singer's Performance:

To evaluate the singer's performance technique proposed by Wei-Ho Tsai and Hsin-Chieh Lee [6], with objective of developing an automatic singing evaluation system for Karaoke performances. This technique has used various acoustic features like pitch, volume and rhythm to assess a singer's performance and improved the singing evaluation capabilities on Karaoke machines. The performance of singer evaluated in terms of technical accuracy then system assign a rating score. It produced the results of automatic singing evaluation such that those are close to the human rating.

- **Advantages:**

- 1) It produces the results of automatic singing evaluation such that those are close to the human rating.

2.4 Feature Extraction Methods:

The main task in music retrieval is feature extraction. One of the approach of feature extraction is given in [7]. The method given to extract features for purpose of key detection. The key detection technique is related to all music related processing such as, music transcription, pitch and time structure discovery, genre classification. Specifically in music theory, a key is the set of musical notes that are primarily used for constructing a piece of music. The key provides important information about the musical content of a music piece, like harmonic and melodic context. For key detection, some methods are already present which are not applicable for key detection from acoustic music because, music notes can't reliably identified.

The general approach for key detection consists of two steps: pitch extraction and key detection. The pitch extraction step outputs the features that defines presence of musical note in particular pitch or pitch class in audio signal. And in key detection step, the keys that are much similar to extracted feature are determined. But this general approach is having issues like mistuning of musical notes and noisy sounds could be removed. So the method proposed in [7] is efficiently used for feature extraction. A tuning pitch determination algorithm has been used for extraction of note partials from signals. This has allowed as much variation as possible. Then system has used consonance filtering for pitch profile feature extraction which used for key detection in music.

Emmanouil Benetos, Simon Dixon, Dimitrios Giannoulis, Holger Kirchhoff, and Anssi Klapuri[8] gives a techniques about the Automatic Music Transcription (AMT). Which is the process of converting an audio recording into some form of musical notation. The core problem for considered in AMT system was the detection of multiple concurrent pitches. Study given in [8] has shown the limitations of current AMT methods and identified directions for future research.

2.5 Feature Selection and Ranking Methods:

For selection of feature which helps in evaluating singer's performance, techniques are required. One of the technique is given in [9]. This method used the feature selection (FS) algorithm for filtering the low efficiency features for fast speech emotion recognition. They have developed an algorithm for feature selection by considering discriminative ability as well as time consumption and redundancy of each feature. Technique given in [9] reduced the dimensionality of qualified features and at the same time improve its discriminative ability for a more efficient and effective emotion recognition system. To retrieve songs matching with singer's vocal competence from database ranking is done. In [10], the given technique describes the method of ranking. In this technique of ranking, given a query, the ranking function assigns a score to each document. Then ranks the documents in descending order of the scores. The ranking order represents the relevance of documents with respect to the query. This is also used for ranking of songs in music retrieval.

3. CONCLUSIONS

The techniques of content based song recommendation are described in the paper are having disadvantages like, they do not consider the ability of singer. So these methods are not efficient for the use of song recommendation. This has given the need of techniques of music retrieval which will consider the singer's ability. Also the paper have described the techniques which will be used in the understanding singer's ability.

ACKNOWLEDGEMENT

This is to acknowledge and thank all the individuals who played defining role in shaping this work. I avail this opportunity to express my deep sense of gratitude and whole hearted thanks to Dr. Dinesh B. Hanchate for giving his valuable guidance, inspiration and encouragement to complete this task.

REFERENCES

- [1] Kuang Mao, Lidan Shou, Ju Fan, Gang Chen, and Mohan S. Kankanhalli, "Competence-Based Song Recommendation: Matching Songs to Ones Singing Skill," in IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 17, NO. 3, MARCH 2015.
- [2] K. Hoashi, K. Matsumoto, and N. Inoue, "Personalization of user profiles for content-based music retrieval based on relevance feedback," in Proc. ACM Multimedia, 2003.
- [3] J. Shen, J. Shepherd, and A. H. H. Ngu, "Towards effective content-based music retrieval with multiple acoustic feature combination," IEEE Trans. Multimedia, vol. 8, no. 6, pp. 1179-1189, Dec. 2006.
- [4] Y. Yu, R. Zimmermann, Y. Wang, and V. Oria, "Scalable content-based music retrieval using chord progression histogram, and tree-structure LSH," IEEE Trans. Multimedia, vol. 15, no. 8, pp. 1969-1981, Dec. 2013.
- [5] K. Mao, J. Fan, L. Shou, G. Chen, and M. S. Kankanhalli, "Song recommendation for social singing community," in Proc. ACM Multimedia, 2014.
- [6] W.-H. Tsai and H.-C. Lee, "Automatic evaluation of karaoke singing based on pitch, volume, and rhythm features," IEEE Trans. Audio, Speech, Lang. Process., vol. 20, no. 4, pp. 1233-1243, May 2012.
- [7] Y. Zhu and M. S. Kankanhalli, "Precise pitch profile feature extraction from musical audio for key detection," IEEE Trans. Multimedia, vol. 8, no. 3, pp. 575-584, Jun. 2006.
- [8] E. Benetos, S. Dixon, D. Giannoulis, H. Kirchhoff, and A. Klapuri, "Automatic music transcription: Breaking the glass ceiling object recognition and segmentation," in Proc. ISMIR, 2012, pp. 379-384.
- [9] L. Zhang, M. Song, N. Li, J. Bu, and C. Chen, "Feature selection for fast speech emotion recognition," in Proc. ACM Multimedia, 2009, pp. 753-756.
- [10] Z. Cao, T. Qin, T.-Y. Liu, M.-F. Tsai, and H. Li, "Learning to rank: From pairwise approach to listwise approach," in Proc. ICML, 2007.