Abstract: With the increasing use of smart devices, many applications have been developed. Listening to music is one of the facilities provided by one of these applications. Users prefer listening to music during their daily chores like while exercising or driving etc. In terms of song searching, user can use voice recognition systems or user can just type the name of the artist or genre or title of the song. But how to search the tune of the song that has been stuck in our head all day long? To solve such problem we are providing QBH (Query by Humming) system where the user can hum the tune and with the help of signal processing, the system will recognize and rank the results obtained based on Personal Hybrid Ranking Method. The effectiveness of the system is characterized by the parameters such as remarkable accuracy, throughput and response time.

Keywords: QBH system, Personal Hybrid Ranking.

I. INTRODUCTION

Searching song by singing/humming section of it is the most natural and easy way to search the song. This search method is the most useful when user don’t have access to audio device or user is unable to recall the attributes of the song like song title, name of artist, name of album. The background of the user often influences the genres of the songs being searched, and this is an area of research seldom studied. In current systems, the recommendation method has not been integrated into the query by singing/ humming systems. We apply this new concept and application. In current query by singing/humming systems, when query results are returned, a user’s preference or previous search history are often not considered. In proposed system, we use the information from a user’s search history, as well as the properties of genres common to users with similar backgrounds, to estimate the genre or style, the current user may be interested in, based on a probability calculation. The accuracy from querying by singing/humming is improved. Our method is divided into two phases. In the first phase, we find the possible search results. This is similar to the conventional singing/humming query process. During the second phase, the musical preference of the user is utilized to rank the possible search results again. Songs that are most likely to be queried would be positioned at the front of the list in the search results. The Major concepts contributing for implementation of system are Genetic algorithm, music database, music recommendation system.

QBH System- Query by humming is a method of searching for a song in multimedia databases that contains the melody description of songs. The database of songs can be searched by hummed queries. The user hums a melody into a microphone that is connected to any handled device and the QBH system searches the database of the songs which are similar to input query and returns the result to the user as a list of songs that match.[5]

Recommendation System- Recommendation method has not been integrated into the query by singing/ humming systems. This new concept and application are applied. In current query by singing/humming systems, when query results are returned, a user preference or previous search history are often not considered.

User Preference- The background of the user often influences the genres of the songs being searched. It uses the information from a user search history, as well as the properties of genres common to users with similar backgrounds, to estimate the genre or style the current user may be interested in based on a probability calculation. The accuracy from querying by singing/humming is improved.

With the advent of new astonishing technologies, consumers can now use internet to play any type of music anywhere, anytime by searching it very easily. Automatic playlist creation, music recommendation or music search are related problems. Searching song by singing/humming section of it is the most natural way to search the song. This search method is the most useful when user don’t have access to
audio device or user is unable to recall the attributes of the song like song title, name of artist, name of album. Therefore ERRQBSH system is intended to implement “Effective Result Ranking of Query by Signing/Humming”, which is the mechanism to search song sung/hummed by user as well as it recommends user by doing effective ranking of the results obtained. Using user’s profile and previous search history and ranking results according is the most challenging task in this system.

II. SYSTEM ARCHITECTURE

The majority of existing QBH systems attempt to improve search accuracy through audio processing or similarity comparison. But, it is very uncommon to calculate retrieval results based on a user’s listening habits or previous search history. We propose to add hybrid recommendations [2] to the search results in the original QBH system. Existing System have implemented query by singing/humming. OBSH system includes The Singing/Humming Signals Process [1]. This process includes conversion of .wav form of user recorded query into MIDI files. Digital music is produced in large quantities and rapidly distributed over the Internet. Methods to quickly retrieve a song have been an important research topic in the area of music information retrieval (MIR). Ghias proposed the first complete system for query by singing/humming. They use a specific string to represent the pitch difference between the melodies of each song in the music library. Prior to comparison being performed between human voices and the music library, descriptions of the melodies are often converted into features or strings. If we code melodies using audio features, similarity calculations and retrieval process are based on Euclidean distance, or cosine similarity. A dynamic programming algorithm [4] is commonly used during the edit distance calculation of two melody sequences in order to save time.

We propose to add recommendations to the search results in the original QBH system. Suppose we have a large library, the recommendation system filters and selects the songs which the user is likely to be interested in. The most common methods currently used by recommendation systems are the content-based filtering, collaborative filtering, and hybrid recommendation methods. Since the content-based filtering method and collaborative filtering method each have their own advantages and disadvantages, some researchers have combined these two recommendation methods into a hybrid recommendation method. The recommended results are then provided on the basis of the user’s personal interests (content-based filtering method) and data items that are read by other similar users (collaborative filtering method). A hybrid recommendation model that can predict a binary user preference was proposed in [2]. The authors aim was to separate items that users vote highly for, from items that users will not vote for. The outcome of this task can be applied on binary user behaviour data.

III. MODULES IN THE PROJECT

- Module recognizing and processing query sung by user.
- Doing basic matching of user query and data in music database
- Re-ranking based on Similar Queries, Ranking by User’s Preference
- Ranking through Similar Users’ Records, Personalized Weight Adjustment Calculation.

IV. METHODOLOGY AND EXPERIMENTS PERFORMED

- The Hybrid recommendation method is used for recommendation.
- The two primary MIR (Music Information Retrieval) methods are used which are key-word based retrieval and content based retrieval.
- The YAAPT method is used to perform the pitch tracking.
- The Edit distance method is used to obtain potential results by considering the songs with small edit distance.
- The Ranking by Preference (RBP) method is used which uses personal query history for ranking.
- The Personal Hybrid Ranking (PHR) method is used that uses a user’s individual information and user’s with similar backgrounds.

DATA FLOW ANALYSIS

Figure 1: System Architecture

Figure 2: Data Flow diagram
The segments sang/hummed by a user are compared with music files in the database. These music files are in an appropriate format in order to perform the comparison. Possible search results are output based on the calculated matching rate. Using this method, many music files from genres that are impossible to be queried, are displayed. In addition, the actual music file required may be positioned further back in the search results, causing inconvenience during operation. And thus as a output we get a list of songs.

V. SIGNAL PROCESSING

We then use a pitch-tracking algorithm to convert the pre-processed audio signal into pitches (fundamental frequencies). This is used to analyze the variations of the melody sang/hummed by the user. The pitch-tracking algorithm can be divided into three methods, namely, time-domain, frequency-domain, and spectral/temporal approaches. For example, the average magnitude difference function (AMDF), the average squared mean difference function (ASMDF) and the other autocorrelation function based methods are time-domain approaches. The popular frequency domain approaches include: the harmonic product spectrum and cepstral analysis. These methods typically utilize the period gram to convert the signal to an estimate of the frequency spectrum. Although the frequency-domain approach provides better accuracy for pitch detection, it requires increased calculation time. The spectral/temporal approach uses time- and frequency-domain approaches to calculate the pitches. This requires the highest level of computational requirements yet also offers good accuracy for pitch detection. A well-known method is known as YAAPT. In the client side of our system, mobile devices with relatively poor calculation power are used. We transfer the signal sang/hummed back to the server via the Internet for processing. The YAAPT method is then used to perform the pitch tracking. A series of fundamental frequencies are obtained after processing the signal sang/hummed using the YAAPT algorithm. To make comparison with the music in the database easier, the frequency information for signals sang/hummed are represented in the form ofnotes. NoteA4meansthefrequency of the note is 440 Hz. Frequencies of other notes are converted based on the number of semitones away from A4. The conversion formula is

\[
F_n = 2^n/12 \times 440 \text{ HZ}
\]

Where is number of semitones away from A4. Notes higher than A4 are positive, and notes lower than A4 are negative. The audio signal sang/hummed cannot constantly stabilize on a particular frequency. The frequencies of each sampling point in the audio signals are matched to the frequencies of then earnest notes. For example, if the input audio frequency is 115 Hz, note A2 is 110 Hz and note is 116.54 Hz, the input signal is classified as. Segmentation of notes is already performed on the input audio signal. After converting every sequence of the input audio signals into notes, the segment with the largest number of notes is chosen as the notes representing this segment.

VI. RESULTS

Our initial motivation for this research was due making difficult to display a large number of search results on the same page. Our work attempts to minimize the need to use multiple pages, so system will first detect successfully, start and end note of segment sang by user. Then it will match segments sang/hummed by the user with the files stored in the database and will perform ranking by similar queries and also ranking by user preference. Finally it will provide an optimized result. This system will give the 76.3% accurate results.

VII. CONCLUSION

The method of querying by singing/humming is the most natural and simple technique to perform music search. An individual does not need to know the title of the song or name of the artist to perform the search. In our study, we use previous search histories of the users as well as users from similar backgrounds to perform probability calculations in order to estimate potential music genres for the user. This improves the rate of success. Such concepts often appear in system design recommendations. We combine the technique with query by singing/humming. Through our experiments, we confirmed that the methods proposed can significantly improve the search accuracy. We used the user search history in the ranking calculation. The search system is hindered by the cold start effect, when a user performs little or no searches in the past, it would cause the accuracy rate to remain at its current level and not improve. In future work, we will study how to reduce the impact of the cold start problem on the system. In this system, it takes a long time to calculate the edit distance between each string. This causes significant delays when returning the results. A user may discontinue use of this system due to the delay. This problem is also an important research topic in order to reduce the time required to calculate the similarity. All of our experiments are currently performed in a quiet environment. In the real world, this is often not the case and it is expected that devices will be operated in noisy environments. A further area of future research is he filtering of noise to reduce the errors during contour string conversion. The existing system requires manual intervention to establish the corresponding MIDI data for the melodies. An automatic system to perform this function also requires study.

REFERENCES

