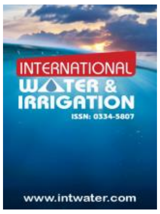

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 01 – 09

Analysis Of Music Multimedia Retrieval and Measurement Quality Model Using Ranking Learning Algorithm

Shanshan Gao[1*]

[1*]The Arts Faculty of North University of China, Taiyuan, Shanxi, 030051, China

*Corresponding author e-mail: gss0703@163.com

Abstract: Reasonable music feature library construction methods are of great significance to the realization of accurate music retrieval. In view of the limitations of current music multimedia retrieval, a ranking learning model is introduced in this paper, constructing corresponding models by fusing local and global indexes, constructing and mapping the attributes and characteristics of music multimedia using feature matrix to clean the data; looking on the global search, nodes are used to search, aiming to improve the efficiency of music multimedia retrieval and measurement quality model construction. The results show that the ranking learning model can effectively support the construction of music multimedia retrieval and measurement quality model.

Keywords: Music Multimedia Retrieval; Global Index; Local Index, Raking Learning Algorithm

1. Introduction

With the continuous development of social economy, there are more and more ways to obtain multimedia. Therefore, in a complex network environment, identifying and searching for the corresponding music multimedia technology from the massive multimedia data has become a difficult point and hot spot in the current research within the industry [1-2]. Traditional music multimedia retrieval is usually difficult to guarantee the efficiency and accuracy, and the retrieval cost is high, therefore, the time cost required is high [3-4]. Based on this, the content-themed music multimedia search is introduced gradually within the industry, but accurate grasp of content and feature extraction are difficult to break through the bottleneck. For example, through the establishment of a music multimedia database, such demand effects become more and more obvious.

The continuous improvement of Internet technology has made it easier to share network resources, and music multimedia resources can also be shared on the Internet. Therefore, traditional search methods such as attribution company, subject, keywords, etc. are far from being able to meet the needs of users. Meanwhile, such retrieval are likely to cause repeated searches, which will lead to excessive search time and relatively simple feature extraction [7-10]. Therefore, in view of these limitations, through the analysis and integration of the global and local retrieval, respectively, the construction and analysis of the music multimedia retrieval and measurement quality model is realized relying on the ranking learning model in this paper,, aiming to improve the efficiency of music multimedia retrieval and measurement.

1.1 Ranking Learning Model Method

Through the investigation of content-based music retrieval schemes by domestic and foreign research institutions, the music retrieval framework system shown in Figure 1 can be obtained. Although the specific solutions of each institution are different, they all include the basic functional units described in the framework. During solving specific problems, most of the attention is focused on pitch extraction, melody contour (music feature) representation methods, and matching algorithms. In these three aspects, there are a large number of papers every year for multi-faceted research, but it has not attracted widespread attention how to form music melody features database. The music feature library is the basis for music retrieval, and only when the ones with perfect feature definitions and correct feature values make the user input more accurately, the correct track can be obtained [11]. If the feature library definition itself cannot fully express the characteristics of the melody

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 01 – 09

change of the music, or if reliable feature values are not generated for various reasons when the feature library is created, it will result in that the corresponding song list output cannot be obtained even if the song feature input is very accurate. It can be seen that a complete and accurate music feature database plays a very important role in achieving accurate retrieval, pitch extraction algorithm and matching algorithm effect testing.

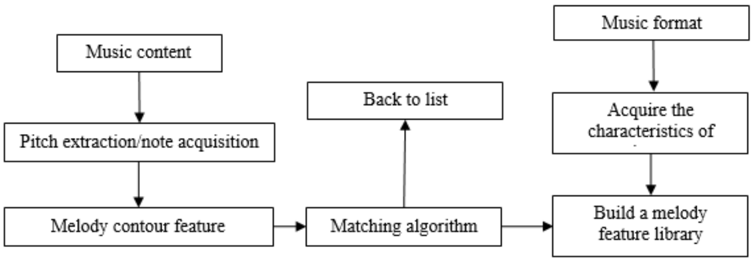

Figure 1: Music Retrieval Framework.

Pitch information reflects the change in the pitch of the song being sung, the pitch information expresses the length of time a certain pitch lasts, and the beat information reflects the speed of humming/singing of a certain song. Based on this This kind of analysis, the definition of the music feature library generally includes these three kinds of information, or only the change information of pitch and duration, the beat information is set as the default rate, and then the elastic matching algorithm is used to achieve the melody feature within a certain beat change limit.

In this paper, the diversity ranking learning framework is used to solve the problem of multimedia diversity retrieval in music and extract a series of features.

The diversity ranking learning framework in music multimedia is shown in Figure 2.

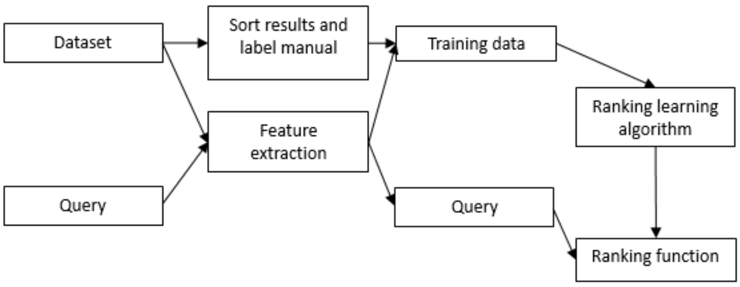

Figure 2: Diversity Ranking Learning Framework.

1.2 Similarity Feature

Based on the foregoing, the core of music multimedia retrieval and measurement is to query the similarity measurement calculation of relevant music resources. Through music multimedia retrieval, the corresponding feature library is established, and the corresponding similarity measurement is established, and is described from three aspects of text characteristics, tag distribution characteristics, social media features, of which the text features are used to describe the text similarity of music multimedia, tag distribution is used to describe the similarity of music multimedia with tags, and social media is used to describe the unique similarity information of music multimedia.

2. Text Features

The calculation method for the characteristics of music multimedia is shown in formula (1).

$$TT_1 = 1 - \frac{t_i \cdot t_j}{|t_i| \cdot |t_j|} \quad (1)$$

Among them, t_i and t_j are the vectorized representations of the blog text respectively.

2.1 Label Distribution Characteristics

Tag distribution retrieval uses the corresponding keyword phrases to cover all music multimedia resources as much as possible. The tag distribution model can be used to detect the implicit subtopics of all blog posts on the same topic and estimate the probability distribution of each blog on it[13-14].

The calculation formula of the topic feature of the tag distribution is shown in formula (2). Figure 3 shows the detected music multimedia tag distribution.

$$TT_2 = \sqrt{\sum_{k=1}^{m} \left(p\left(z_k \middle| t_i\right) - p\left(z_k \middle| t_j\right) \right)^2} \quad (2)$$

Among them, $p\left(z_k \middle| t_i\right)$ refers to the probability of the blog $tweet_i$ on the subtopic z_k.

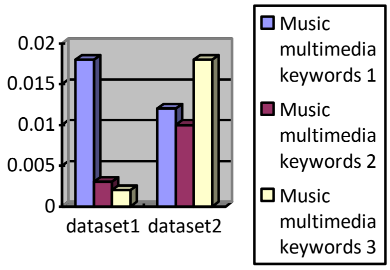

Figure 3: Probability Distribution of Sub-Topics And Blog Posts On Sub-Topics.

2.2 Social Media Features

Music multimedia has certain characteristics of network social multimedia [15-16].

1) Time characteristics

The calculation of the time feature is based on two time-normalized timestamps (normalized by the maximum value and the minimum value), and the calculation method is shown in formula (3).

$$TT_3 = \left| t_{norm}\left(t_i\right) - t_{norm}\left(t_j\right) \right| \quad (3)$$

In the formula, $t_{norm}\left(t_i\right)$ and $t_{norm}\left(t_j\right)$ represent the normalized representation of the music multimedia release time, respectively.

2) Topic features

The similarity of the subject terms involved in two music multimedia resources are calculated through the correlation index. The specific calculation is shown in formula (4):

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 01 – 09

$$TT_4 = 1 - \frac{\left|Term(t_i) \cap Term(t_j)\right|}{\left|Term(t_i) \cup Term(t_j)\right|} \quad (4)$$

For the current research on the retrieval and measurement of music multimedia, the following steps are usually adopted for retrieval and measurement, as follows:

1) The music audio signal is processed;

2) A query index of resources is constructed by extracting features;

3)A database audio index is built;

4) A query mechanism is built;

5) According to the similarity of corresponding music multimedia search results, retrieval and measurement of different types of music resources are performed in different categories.

In the actual retrieval process of music multimedia, the music multimedia formats usually involved are.mp3,.wav,.midi and other formats.

At the same time, at the core of actual music multimedia retrieval, the feature-based matching algorithm is shown in Figure 4:

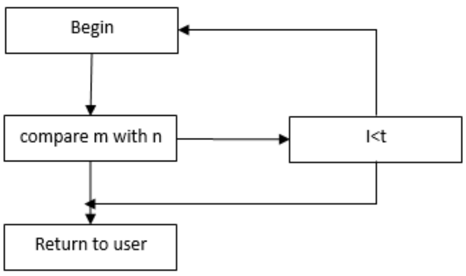

Figure 4: Feature String Matching Algorithm.

According to the summary and condensing of the above retrieval methods, there may be some fixed attributes. Therefore, this requires significant repetitiveness in music multimedia retrieval. Meanwhile, in order to improve the search efficiency in the music multimedia library in the shared network, a parallel decomposition algorithm can be used to further improve the resource search efficiency of the music multimedia under the sharing efficiency. Therefore, a comprehensive model combining local search and global search is proposed in this paper. The specific process and ideas are shown in Figure 5:

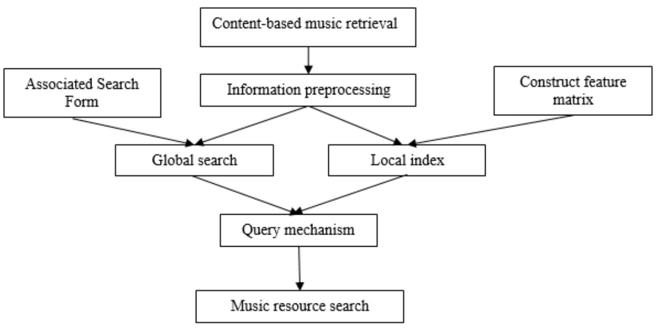

Figure 5: Design Ideas of Music Retrieval Model.

3. Construction of Model

3.1 Description of Node Relationship in Unstructured Network

Compared with traditional text information retrieval, music files need a more efficient retrieval algorithm. Therefore, in order to improve the effect of indexing, a special index structure is usually constructed for music files. However, the construction of the music file index structure requires a larger storage space and a higher-performance query capability. Traditionally, the construction of music index is done by dividing music fragments. A method is proposed in this paper that does not need to consider the length of traditional music fragments, but directly constructs a global index and a local index among each node. A global index and a local index are constructed in every node, and users at different nodes can search for resources in this range.

3.2 Information Pre-processing

Among music files of different formats, there will be a large number of repeated segments in each song, which greatly increases the complexity of the index structure. To improve the efficiency and accuracy of the index, the first step is to solve the repeated fragments, music file information must be ed to eliminate redundancy. Therefore, in order to improve the accuracy, all the topic information is extracted and the redundant information is removed in this paper. Through such a method, it cannot only reduce the overall complexity, but also effectively guarantee its accuracy.

In music, pitch is usually more able to influence the theme of a piece of music. Therefore, the theme and standard deviation are used as the two key theme feature factors in this paper, assuming that it is the pitch reflected by the i-th note in a musical melody, its average pitch and standard deviation are calculated as shown in formula (5) and formula (6)

$$\overline{p} = \frac{1}{k}\sum_{i=1}^{k} p_i \tag{5}$$

$$\overline{d} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(p_i - \overline{p}\right)^2} \tag{6}$$

In formulas (5) and (6), k is expressed as the number of notes in a certain theme.

Through the treatment of the subject, the $\overline{p}$ and $\overline{d}$ of two different themes are obtained. In this way, all the themes in the folk music song are extracted, and the $I = I_1, I_2 \cdots, I_n$ of themes can be obtained respectively. After re-calculation, different I values are obtained, thereby eliminating the repeated fragments, reducing the fragments of retrieval, and greatly shortening the time of retrieval.

3.3 Query Mechanism Construction

In the process of building a music feature library, the pitch detection technology in audio signal analysis can be used to achieve music melody acquisition and build a library. Specifically, it includes 3 steps: monophonic wave file acquisition, melody automatic segmentation, and feature storage.

3.4 Monophonic wave File Acquisition

The pitch extraction method processes time-domain signals, and the wave format file is required for the sound signal to be processed, or other compression formats are converted to wave format. Monophony is another requirement for wave format files, that is, the wave file used to build the library should only contain the performance of the main melody by one instrument. The song usually includes a vocal singing part and a background music accompaniment part. The main melody of the song should be defined as the vocal singing

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 01 – 09

part of the original song. Because the vocal input to achieve retrieval is generally for the part of the melody with lyrics, so the best A wave file is a recording of a musical instrument's performance of the main melody. But it is not an easy task to collect this kind of digital music with special requirements. A good solution is to first obtain the monophonic midi files of the songs to be stored in the library, and then use the "MIDI2WAVE" software or internal recording to achieve the generation of the target wave file.

3.5 Acquisition of Music feature

(1) Pitch Detection Method

After obtaining the appropriate wave file, it is necessary to extract the pitch of it to obtain the characteristics of the melody change. The autocorrelation function (ACF) method is widely used because of its simple implementation and good detection effect.

Due to the interference of environment, current and other noises, there are a large number of deviation points in the obtained pitch sequence. The traditional smoothed filter algorithm will remove the grain deviation points while affecting the inherent high and low mutations of the original waveform. Therefore, in order to solve this problem, a " small cabinet " method is proposed, which can well achieve the purpose of filtering out deviation points and retaining inherent mutations. The specific steps are as follows:

The pitch sequence is denoted as X(m), the filter width is defined as L, and the fluctuation range of the pitch value is Range, a two-dimensional array A [Range, L] is defined, and Y(m) to save the filtered data. The implementation steps are as follows :

Step 1. Create a two-dimensional array A[Range, L], take the filter width as its abscissa, and the pitch value fluctuation range as its ordinate.

Step 2. Apply for an array space as large as the pitch sequence array as the storage space for the values after filtering.

Step 3 Intercept the first L pitch values of the pitch sequence and put them into the array A according to the size of the value. Each column in A represents a small cabinet.

Step 4. Count the number of data stored in each small cabinet in array A, find the column with the largest number of data, and calculate its average.

Step 5. Write the mean value as the pitch value after filtering into the corresponding position of the new pitch sequence space.

Step 6. Clear the two-dimensional array A.

Step 7. Move the filter pointer backward one position.

Step 8 Repeat the operation in Step 3~Step 7 for the data whose pitch order is between $\left(\left\lceil \frac{L}{2} \right\rceil, \left\lceil m-\frac{L}{2} \right\rceil\right)$, until the array data is exhausted.

The above method is the situation after completing a filtering. The experimental results show that multiple filtering can keep the waveform change stable and achieve the best filtering effect. After completing this step, you can obtain regular information that can truly reflect the pitch change of the query segment.

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 01 – 09

(2) Automatic segmentation of melody

During the construction of the music multimedia feature library, if the automatic segmentation link is not considered, it will lead to the matching of the user's feature query with the song, so that the feature sequence can be matched with incomplete songs, which reduces the accuracy of the matching algorithm and improves the corresponding time cost.

3.6 Feature Storage

After obtaining the music melody change information and sentence segmentation information, the storage can be realized in the following three forms:

(1) Direct storage. the pitch sequence with segmentation marks is stored directly in a binary file, a file is built for each piece of music, and read in order to achieve retrieval when matching is achieved. This method is simple to implement, but when the number of music pieces is large, all files need to be loaded in sequence for retrieval, which is time-consuming.

(2) Establish a specific file structure storage. In this research, Shanghai Jiaotong University established a custom ref file to realize storage in the form of relative pitch difference and duration.

(3) Store in a vector format. That is, store a series of pitch sequences in the form of feature vectors (pitch, duration, breakpoint or not).

The advantages of the latter two methods are saving storage space and faster loading during retrieval. The disadvantage is that there are strict requirements for the post-processing of the pitch sequence, and the pitch value must be strictly equal in the steady sounding segment of each note.

One of the most direct ways to obtain music melody change information is to read the numbered musical notation or staff of the music. It is also easy to collect the scores of a large number of songs in a short period of time. There are two ways to build a library with score information:

(1) Directly manually enter the score information. Enter the note information, beat information, and segmentation information of the music, and save it directly in the database. This method can achieve segmentation due to manual participation, but the efficiency is very low, and it can only be used when building a large-scale database, and it is also prone to omissions due to manual intervention, and it is not easy to detect.

(2) Optical score recognition is used to extract note, duration, beat and segmentation information. The highest recognition rate can reach more than 90%.. Through the OMR technology, the score information can be converted into MIDI files, which also provides convenience for building a library through MIDI files.

4. Experimental Verification

A P2P network is constructed through the PeerSim simulator, which contains 2000 nodes, and the number of connected nodes around each node is kept at 1-10, and 1000 ethnic songs are prepared simultaneously. Through the extraction of these 1000 ethnic music, and a random way is used to put them into different nodes. Finally, the results shown in Figure 6 and Figure 7 are obtained through inspection.

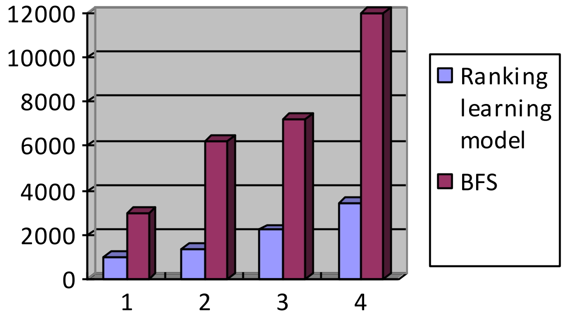

Figure 6: Response Time After Partial Index Establishment.

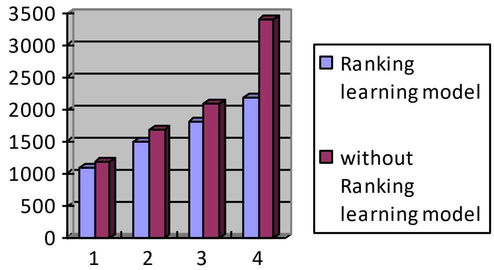

Figure 7: Changes In the Number Of Queries Under The Global Index.

5. Conclusion

Based on the ranking learning model, the construction and analysis of the music multimedia retrieval and measurement quality model are realized through the analysis and fusion of the global and local retrievals, the retrieval efficiency of the music multimedia is improved, and a good reference mode for improving the retrieval efficiency of the music retrieval system is provided. Simulation experiments prove the effectiveness of this method.

Data Availability: The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

Funding Statement: This study did not receive any funding in any form.

References

[1] Rossetto L , Giangreco I , Gasser R , et al. Open-source column: content-based multimedia retrieval using vitrivr[J]. ACM SIGMultimedia Records, 2018, 9(3):8-18.
[2] Yunping, Wang, Xin, et al. Performance Analysis of Multimedia Retrieval Workloads Running on Multicores.[J]. IEEE Transactions on Parallel & Distributed Systems, 2016, 27(11):3323-3337.
[3] Fadel S G , Torres R . Neural relational inference for disaster multimedia retrieval[J]. Multimedia Tools and Applications, 2020, 79(35):26735-26746.
[4] Lu Y , Sebe N , Hytnen R , et al. Personalization in multimedia retrieval[J]. Multimedia Tools & Applications, 2015, 51(1):247-277.
[5] Purswani N , Ramrakhyani R , Makhija M . Content based Multimedia Retrieval using Automatic Speech Recognition[J]. International Journal of Computer Applications, 2015, 118(2):7-9.
[6] Schedl, Markus. Investigating country-specific music preferences and music recommendation algorithms with the LFM-1b dataset[J]. International Journal of Multimedia Information Retrieval, 2017, 6(1):71-84.
[7] Alqaraleh S , Ramadan O . Utilizing Query by Example for Fast and Accurate Multimedia Retrieval[J]. Applied Mathematics & Information Sciences, 2015, 9(1):125-134.
[8] Dias R , Fonseca M J , Silva N , et al. EnContRA: a generic multimedia information retrieval meta-framework[J]. Multimedia Tools and Applications, 2015, 74(11):3691-3713.
[9] Furtado S D , Michael C C , Yan Z , et al. Fast Similarity Matrix Profile for Music Analysis and Exploration[J]. IEEE Transactions on Multimedia, 2018, 3(1):1-10.
[10] Papadopoulos S , Cesar P , DA Shamma, et al. Social Multimedia and Storytelling [Guest editors' introduction][J]. IEEE Multimedia, 2015, 22(3):10-13.

[11] Clerico A , Tiwari A , Gupta R , et al. Electroencephalography Amplitude Modulation Analysis for Automated Affective Tagging of Music Video Clips[J]. Frontiers in Computational Neuroscience, 2017, 11(2):115-123.
[12] Macfarlane A . Knowledge Organisation and its Role in Multimedia Information Retrieval[J]. Knowledge Organization, 2016, 43(3):180-183.
[13] Bouafif L , Ellouze N . Speech-Music-Noise Discrimination in Sound Indexing of Multimedia Documents[J]. Computers, Materials and Continua, 2019, 61(3):1-10.
[14] Lampropoulos A S , Tsihrintzis G A . Content Description of Multimedia Data[J]. Springer International Publishing, 2015, 10 (4):63-76.
[15] Gupta R , Abadi M K , JA Cárdenes Cabré, et al. A Quality Adaptive Multimodal Affect Recognition System for User-Centric Multimedia Indexing[J]. ACM, 2016,5(2):317-320.
[16] Fang J T , Day C T , Chang P C . Deep feature learning for cover song identification[J]. Multimedia Tools & Applications, 2017, 76(22):23225-23238.