Survey of Content-based Music Information Retrieval

DENG Jie

Abstract

As the digital music becomes more and more popular and huge, there are a wide range of music services, especially in music information retrieval (MIR). With the rapid progress of the information science and technology, the content-based music analysis and processing become hot area. MIR seeks to utilize the theory of music to better search database of music contents through musical expressed queries, that is to say, searching music by music. This area research has significant commercial and research promise in nowadays.

This paper mainly gives a comprehensive review of the latest year's research on MIR. In addition, some key techniques and approaches to MIR will also be introduced, particularly focusing on content-based MIR. Some main issues of content-based MIR will also be summarized. Moreover, the paper also introduces some basic music terms, concepts and music processing basics. At the same time, some high level abstraction such as semantic description has broadened the future of MIR.

***Key words**: music information retrieval, content-based*

1. Introduction

With the fast development of the Internet and digital devices, there are enormous amount of digital music available for people accessing and downloading via PC and mobile device. People have great music information needs, for example seeking music songs accurately and finding similar songs. In addition, music information retrieval has a wide rang of potential applications. As the emergence of social network, music reviews, ratings, and some similar or personalized recommendations are becoming more important [1]. Therefore, they all bring great challenge to music information retrieval. Currently, music information retrieval is becoming a hot research area for solving the users' particular interest and needs. As the music is an art form whose elements are pitch, rhythm, dynamics, timbre, texture and so on [2], it's very difficult to describe them exactly. In addition, music is relevant to human auditory perception, which expresses emotion and mood. It is also very difficult to measure them accurately. Therefore, at present, most of the music search engines use some metadata (the name of songs, artist and album, lyrics, etc.) to classify and retrieval music, which is low-level search conditions. Taking Google and Yahoo as an example, they all use text-based metadata to search music. However, content-based MIR searches music by the conditions of musical terms and concepts such as melody or harmony to describe the content of music, which is a high-level search using real musical features. Because there have less efficient and effectiveness approaches to content-based MIR, less search engines use this retrieval type, such as Midomi.com, and ThemeFinder. However, the improvement of the retrieval approaches and technologies are also made constructively in this area. Currently, most of the researches on music information retrieval focus on the content-based, and the main idea on this approach are describing a set of music features computing directly from their contents [3]. Although there are some problems, researchers have already developed some systems to allow users retrieval music based on their content descriptors. MIR shares many problems of traditional information retrieval. In addition, there are also mainly three problems in the MIR: music feature representation, music queries and music feature matching. Most of current content-based MIR systems mainly use melody information, which consists of a combination of music notes and pitch. Music queries have vivid ways, for example users can query by humming or query by using keyboard imputing key words or some professional music notes. Matching techniques are also key retrieve step, which adopts different similarity measures to quantify the similar music works. There have already some content-based MIR approaches on melody feature matching.

This paper mainly focuses on the review of techniques and approaches for content-based music information retrieval. The paper is organized as follows. In section 2, a brief review of music theory will be given, which consists of basic music concepts and characteristics. Section 3 gives a detailed description on music information retrieval in real world, which introduces the user intent, music data and format, music query mode and processing, and some existing music search systems. Current music information retrieval key techniques will be explained in the section 4, followed by evaluation strategies of the MIR, which are discussed in Section 5. Finally, some conclusions and future work are given in Section 6.

2. A Brief Review of Music

This section will briefly introduce some music theory which is related to MIR. Music information retrieval is based on three-filed subjects: traditional information retrieval, musicology, and digital audio. The following figure 1 shows the content of computer music.

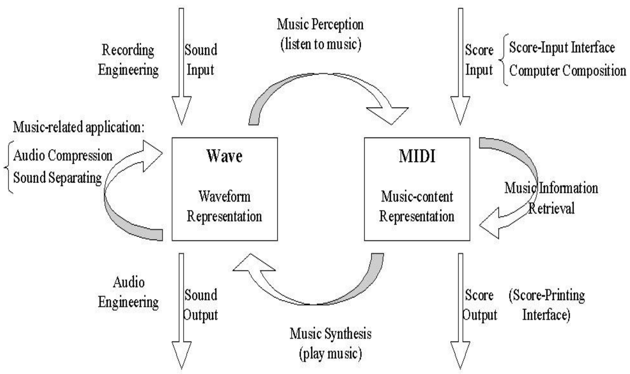


Figure 1, Computer Music

The following will give a brief introduction to music theory. Nicola Orio [4] has already given a review of music concepts and characteristics in "Music Retrieval: A Tutorial and Review".

2.1. Music Concepts

Music is an art form, which produces sounds and silences in time. Based on the music theory [5], there are three basic musical sound elements: pitch, dynamics, and timbre. Pitch is the range from low to high, which dependent on frequency of the sounds. Dynamics is the intensity of the sounds, which is related to the sound amplitude. Timbre is the quality of a music note or sound that distinguishes different types of sound production. Apart from these three basic music elements, there are also many other common terms used to describe music [4], for example tempo, tonality, time signature(how many beats per measure) and key signature(symbols # and b). In the real world, the staff is the basis of written music. Music note is represented symbols to express the length of the pitch. The following figure 2 shows polyphonic musical score measures.

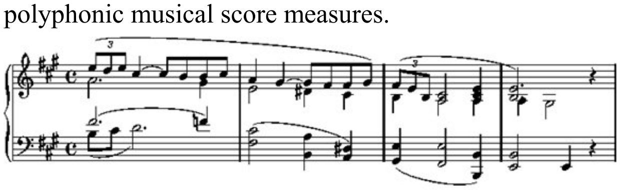


Figure 2, Example of a musical score (excerpt from Premi`ere Arabesque by Claude Debussy).

2.2. Music Characteristics

Commonly, music has two dimensions [4]: horizontal and vertical. In vertical dimension, for polyphonic music, the two or more sounds are playing at the same time, which results in a relevant perception and simultaneously active and aligned in the musical score. In horizontal dimension, it associates time to the horizontal axis. In addition, there are also other main dimensions of music which could be effectively used for music retrieval, which consist of timbre, orchestration, acoustics, rhythm, melody, harmony and structure. Therefore, in theory, any of the above dimension or their combinations is able to use for a relevant descriptor of a music work.

However, in practice, the most used dimensions are melody, rhythm, and timbre. As the melody is the soul of music, which is an organized succession of pitches, and it is the foundation of the music. In addition, melody can be represented by the main theme of music works. Thus, people always can easy memory music through melody. Rhythm is another key characteristic and will give us intuitive feeling, which stands for the movement of the music in time. Timbre is the quality of musical note which is used to distinguish the different music instruments. For the content of music not just the composer and his or her audience, but the occasion for which it was written and the time and place in which it was performed, it is not easy to express like pictures, because it is just a sense for human beings. Thus for the effective and efficient music retrieval, the above terms and characteristics in musicology are very useful in MIR.

3. MIR in the Real World

As the great use of digital audio, people are able to enjoy and spread music conveniently. One can generate amount of digital music with different styles and formats. Music information retrieval today exists with extremely diverse characteristics and content and different users. Thus, all these factors have inspired many companies such as Google, Yahoo attempting to provide convenient, effective and efficient music search services. Therefore, music information retrieval is now a popular research area and has a wide range of applications for both consumers and musicians.

As far as technological advances are improved, content-based music information retrieval has grown rapidly [15]. However, real world applications of the music retrieval technology are currently limited [10].

3.1. Users and Search Mode

In Internet, there are a number of users interact with music search engines to retrieve what they need. Commonly, there are three-category users: casual users, professional users and music theorists or musicologists. Based on the different music knowledge levels, there are different search modes for different level users.

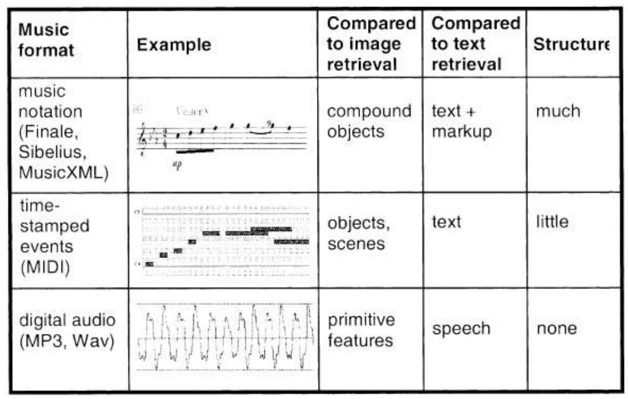
The most music search systems (e.g., iLike, Lala, Pandora.com) are based on the texts or tags, which are rich and expressive to describe the music. For example, these texts may contain artist, albums, tracks, lyrics, genre, and comments, etc [22]. Thus, the users search music only by inputting some of these keywords. In addition, some commercial music website may use these texts or tags to mine and estimate song similarity and artist similarity. Then it retrieves songs that the users are favorite and satisfied.

However, a fewer music search systems (e.g., midomi, tunespotting) are based on the sound or voice, that is to say they search music by singing or humming or playing part of a song to the computer or digital devices. The longer you sing or hum or play the more accurate the results will be retrieved for you.

Last but not least, fewer music search systems are suitable for musical professionals. Some search engines (e.g. Themefinder) identify common themes in Western classical music, folksongs and Renaissance, which require you to have some musical knowledge. When searching music, you need to input some related music parameters, for example, pitch, interval, scale degree, gross contour and refined contour. The others (e.g., tunespotting) are based on the music notation, for example stave. The search engines use the Parsons Code, a rough description of the melodic contour to search music.

3.2. Music Data and Format

Understanding the music data and scope and format plays an important role in the complexity of music retrieval system design. The music data in this paper mainly refers to digital music audio data. Music is usually divided into three categories based on the amount of concurrency: Monophonic, Homophonic, and Polyphonic [8]. Usually, digital audio recordings may be labeled with metadata. Metadata can be used to name, describe, catalog and indicate ownership or copyright for a digital audio file, as well as allow user characterizations of the audio content (ratings, tags, and other auxiliary metadata). When audio formats moved from analog to digital, it maybe embeds these metadata with the digital content itself. There are mainly three musical formats: Symbolic format, Audio format and MIDI [9]. However, in computer, all the music is digital information, which has two expressions: based-on musical instrument digital interface and waveform data, for example mp3, wav. The most uncompressed audio file format is based on pulse code modulation, which samples analog signal, then quantize, and encoding to gain digital music files. MIDI is an industry standard which enables electronic musical instrument to communicate with computers. Figure 3 shows a comparison of current common music formats.



Music format	Example	Compared to image retrieval	Compared to text retrieval	Structure
music notation (Finale, Sibelius, MusicXML)		compound objects	text + markup	much
time-stamped events (MIDI)		objects, scenes	text	little
digital audio (MP3, Wav)		primitive features	speech	none

Figure 3, Common Music Formats

3.2. Music Processing

The first step of music information retrieval should be able to automatically process the musical file to extract useful features or content descriptors. Different musical file formats should adopt different extracting methods. Early research [17] mainly focuses on symbolic music, for example SMF and MIDI, due to it can easily extract melody features and mp3 and wma audio formats are not popular at that time. Currently, most of research on music feature extraction focus on the digital audio form (mp3). As the most important music feature is melody, extraction and segmentation of the melody are the key steps for music retrieval based on the melodic information.

In symbolic form, it’s very easy to extract melody in monophonic score, while it’s more complex to extract melody in polyphonic score, but the most common using statistical approaches. As the melody can be view as a sequence of hundreds of symbols in the symbolic form, thus shot sequence of the melody are more efficient for melody search. N-grams are often used to segmentation. In addition a priori music knowledge will also be used to segment melody. In audio form, the extraction and segmentation is more complex. As the pitch and length of music note are the major elements of melody, the important process in audio music processing is extracting the pitch and determines length. The following sections will give a detailed description to these features extraction and retrieval process.

3.4. Overview of Some Existing Music Search Systems

There are some common functions in some existing music search engines [6].

- Search by music related metadata: (artists, albums, tracks, music reviews, new release, etc.) Yahoo! Music and Allmusic are the examples of this search type.
- Search by music lyrics: Lyrics.com and SongLyrics.com
- Music Media Management and Track Identification: Identify metadata for music tracks, for example Gracenote and MusicIP.
- Recommend similar music: by mining some music feature elements (melody, rhythm, tone color, etc) to recommend user some similar music.
- Recommend personalized music: by mining some users' information to recommend them some their favorite music.

In addition, Musipedia can search music by whistling a theme, playing it on a virtual piano keyboard, tapping the rhythm on the computer keyboard, or entering the Parsons code. It also offers three ways of searching: Based on the melodic contour, based on pitches and onset times, and based on the rhythm. In high level, some system may cross media retrieve supporting natural language queries like mood, which contains semantic information [11].

4. MIR Key Techniques

As the past ten years research on music information retrieval, there have developed some key techniques and methods on music representation, music feature extraction, segmentation, clustering and music matching. Tseng [12] has given a content-based model for music collections. The following part will focus on the content-based music information retrieval key processes and steps. Figure 4 shows the framework of the content-base MIR. Michael Fingerhut presents the global MIR map [9], which contains three hierarchies: music layer (author, composer, performer, etc.), stored data layer (symbolic, text, sound) and information layer (music acoustical, theory, and musicology).

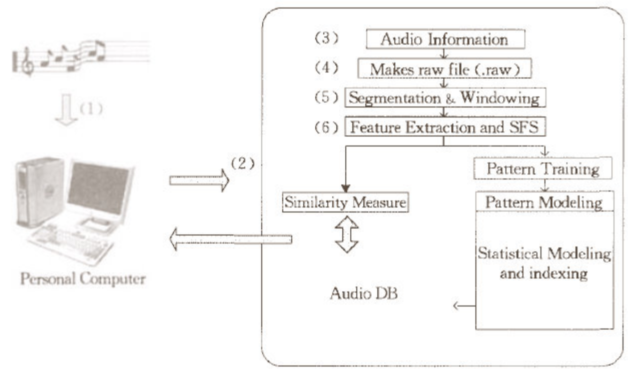


Figure 4, Framework of Content-based MIR

4.1. Basic Approaches to MIR Analysis

In digital audio [7], the recordings are sound waveform, which expresses the frequency changes. In all the sound, the most important element is pitch. A sound consists of a fundamental frequency and a set of harmonic frequency. Thus a sound can be expressed using the following formulation.

$$S (f, t) = \Sigma amp (k, t) \sin (kft + phase (k))$$

Here n stands for number of frequency, amp (k, t) stands for the amplitude of frequency k in time t. In the representation of music, every music note stands for a special frequency. When analyzing the mp3 audio format, using the relationship of fundamental frequency and harmonic frequency is necessary.

When extracting audio features, we should first window audio file, which divides it to small segments, and every segment is called a frame, then use Fourier transform or fast Fourier transform, and finally compute the audio feature values. There are two categories features: perceptual feature (pitch, dynamic, rhythm, etc) and non-perceptual feature (Mel-Frequency Cepstral Coefficients, Average Zero-crossing Rate, and Fundamental Frequency, etc). MFCC is a feature set popular used in speech processing and music modeling.

Audio music segmentation is implemented by detecting the features such as fundamental frequency or energy to segment edge. The edge is the suddenly change point of the audio features. By using automatic audio music segmentation, it's very convenient to segment the accompanying and sing part. There are some melodic segmentation approaches and techniques [18]. Fixed-Length Segmentation extracts from a melody of subsequences of N music notes. In addition, Data-Driven Segmentation can be considered as an extension of N-grams approach. Perception-Based Segmentation will also segment melody according to theory on human perception. Mixed Segmentation and Data Fusion will describe the melody with a set of note sequences. Query Segmentation requires will apply the same segmentation to music files and queries.

When combination of different audio music features has done, it can be used to index audio music files by feature vector. Therefore, different similarity algorithms are able to adopt by computing these feature vectors distance. Then the similar audio music file will be clustered automatically. Therefore, when users query by an audio music sample, the system will automatically extract the feature and then formulate the query vector. Subsequently, the distance between query vector and vector space or previous formulated music cluster vector will be calculated. The threshold will be defined to determine the similarity criteria. In addition, the similarity value can be also used as the output result rating.

4.2. Music Representation

From the music theory, music melody and rhythm are the most used content-based MIR conditions. In fact, most music features used to represent music are always melody and rhythm. In fact, there are mainly three basic melody representations: absolute, relative, and melody contour. Absolute method uses characters or numbers to encode absolute pitch and length, while relative method uses music interval as the pitch vector. Melody contour method [5] uses three characters to express the basic contour of the melody. Figure 5 shows the pitch contour descriptor and gives an example.

$$PC_k(Pitch_{k,i}) = \begin{cases} "U" & if(Pitch_{k,i} > Pitch_{k,i-1}) \\ "S" & if(Pitch_{k,i} = Pitch_{k,i-1}) \\ "D" & if(Pitch_{k,i} < Pitch_{k,i-1}) \end{cases}$$

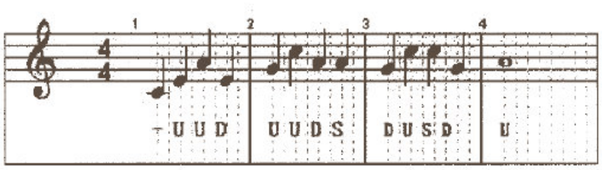


Figure 5, Melody Contour

In this method, S stands for the same with the previous note. U stands for higher than the previous note, while D stands for lower than the previous note. Comparison to the absolute and relative methods, contour can not retrieve the result accurately, but it decreases the search vector space. Rhythm feature representation only considers the rhythm, which omits the melody. If it utilizes character to represent different length of music note, maybe a short music score requires much encoding characters, which is not efficient and convenient. In addition, as timbre is related to the spectrum of sound, Fourier transform is the most frequently used tools in timbre representation and analysis.

4.3. Music Feature Extraction

The majority of music features are melody, timbre, rhythm, thus this part will mainly explain extracting these features [25]. As to melody extraction, pitch is the main element. There are two-category algorithms: time domain and frequency domain models. Autocorrelation function (ACF), Average magnitude difference function (AMDF), and Simple inverse filter tracking (SIFT) are all common used time domain methods. Harmonic product spectrum and Cepstrum are frequency domain method.

Rhythm is the most useful features for extraction. Most popular music and rock music always use tempo tracking to recognize rhythm. It's very difficult to describe timbre, but it has a very close relationship with performance instruments. Beside, Mel-Frequency Cepstral Coefficients (MFCC) has already used as a content descriptor of musical sound. Short-Term Fourier Transform Features (FFT) [20] is a set of features related to timbre textures and not captured using MFCC.

4.4. Music Retrieval

There is mainly two-category content-based music information retrieval: symbolic data and audio data. Researchers have already made a lot of research on MIDI, which transmits pitch and intensity of musical notes to play. However, waveform music data has widely used nowadays, thus waveform music retrieval becomes hot. The following part will introduce these two category music information retrieval approaches.

4.4.1. Searching Symbolic Data

Melodic retrieval based on sequence matching [14] is to retrieve music by strings. Strings are able to represent melody, rhythm, and contour and interval sequence and so on. Thus the most of string matching techniques have been applied in using pitch and melody contour. Melody contour has used this approach, which is represented by only three character symbols. Another approach applies pattern matching techniques to query the sequence.

Polyphony music retrieval usually adopts n-grams model to index and match music notes. In this model, every music symbols will be converting to sequence of melody and interval. It highlights musically relevant sequences of music notes, which will undergo a number of different normalization. Patterns were computed using pith or rhythm or their combined information. Alternatively string matching, statistical model based methods compares the query probability of attribute to corresponding attribute. Most of the current systems use hidden Markov Mode (HMM) to model a set of themes which are extracted from musical files. In addition, a combined method whose distance function is dynamic time warping (DTW) has been computed.

Sequence matching may be very efficient, but it may require the sequence compared to the entire music file in music collections. In addition, some geometric approaches have been applied to find the best matches with the geometric representation of each music file.

4.4.2. Searching Audio Data

Waveform music file is different from audio symbolic music, which is not able to structure the music directly. The typical waveform music retrieval system contains two parts: building music database and query process [13]. Building music database contains feature extraction and computation for example energy distribution and contour. There is a great effort to extract features of every music file in collections. Currently, music fingerprint is the hot technology to analyze sound wave of music files, which is able to represent a piece of music. Similar or the same music maybe has the same music fingerprint. The process of music fingerprint extraction consists of preprocess, Short-time Fourier Transform. Then it will retain unique identification code, which are also used to recognize music file.

For the high level of search, perception feature based method is proposed to retrieve audio music files. This method extracts the abstract description in audio files, then it will window the audio waveform, and then use Discrete Fourier Transform (DFT) to get parameters and frequency energy, and finally compute the relevant feature vectors value. At last, the similarity will be computed from the query audio to all the music files in database. Lie Lu [23] has proposed a new approach to query by humming. Forevermore, collection-based method has been proposed, which use feature extractor to convert PCM signal to collection form, making processing them like processing music note collections. In addition, Self-organized Map technology is also adopted to retrieve audio music files.

4.4.3. Similarity Measures

When the system retrieves the music file, the search result is also very important. Similarity measures will contribute to the search result. From the above music retrieval key techniques introduced, the most of the music similarity measures are based on audio features, for example MFCC, and computed with statistical approaches, for example the Gaussian Mixture Models (GMM). There are also some popular distance measures applied in similarity computation in MIR, such as Euclidean, Hausdorff, Mallows, IRM and so on.

As the social network fast development, some social similarity measures are also becoming important. They are based on the tags, reviews, ratings, etc, which are able to discover the relevant information, while not detecting the audio music files. In addition, the combination of these two ways may have better search results to users.

5. Evaluation of Music Information Retrieval

The evaluation of music information retrieval system is necessary for the system's effectiveness. Symbolic files (MIDI, etc.) and audio files (mp3, wma, etc.) are all should be evaluated. Downie and Nelson [20] have proposed an effectiveness evaluation of n-grams, which is using statistical analysis. There are mainly two popular evaluation measures [19]: precision and recall. Precision refers to the percentage of the retrieved music files which are relevant to the query conditions. Recall refers to the percentage of all the relevant music files in the retrieve music database.

$$\text{Precision} = \frac{\text{Number of retrieved music that are relevant}}{\text{Number of music collections retrieved}}$$

$$\text{Recall} = \frac{\text{Number of retrieved music that are relevant}}{\text{Number of relevant music}}$$

In 2004, Audio Description Contest first attempted to build comparative benchmark of MIR algorithms, which has five different tasks: genre classification, artist identification, tempo induction, rhythm classification and melody extraction [27]. It also provides some training and test data. In the following years, it has added other tasks: music similarity and cover song identification. J.Stephen Downie has already described the scientific evaluation of music information retrieval systems foundations and future [29]. Major issues have been proposed to address the complex nature of music information, adequately capturing the complex nature of music queries, recognition of MIR "relevance" problems.

6. Conclusions and Future Work

This paper gives a comprehensive review of music information retrieval. A brief review of music and some basic music theory are also presented. There are three category users: casual user, professional users and music theorists or musicologists. Based on the different music levels, there are different search modes for different level users. Most music feature used to represent music is also melody, rhythm and timbre. There are mainly three melody representations: absolute, relative, and melody contour, which uses U, D and S these three characters. For melody feature extraction, time domain and frequency domain models are applied. ACF and SIFT are separately these two methods. Tempo tracking usually used to recognize and extract rhythm. FFT is a set of features related to timbre textures and not captured using MFCC. Symbolic data (MIDI) search and audio data (MP3, WMA) search are different content-based music information retrieval category. The paper also introduces some evaluations of MIR, Music Information Retrieval Exchange (MIREX) evaluates frameworks and test music collections. For future work, high level semantic music information [28], for example, emotion, mood, and music content abstract will be expanded and settled.

References

[1] Jin Ha Lee and J. Stephen Downie. Survey Of Music Information Needs, Uses, And Seeking Behaviours: Preliminary Findings. In Proceedings of the Fifth International Conference on Music Information Retrieval: ISMIR 2004, Barcelona, Spain, 441-446, 2004

[2] J. Sephen Downie. Music Information Retrieval (Chapter 7). In Annual Review of Information Science and Technology 37, ed, 2003

[3] Michael S. LEW, Nicu Sebe. Content-Based Music Information Retrieval: Current Directions and Feature Challenges. Proceedings of the IEEE, April 2008

[4] Nicola Orio. Music Retrieval: A Tutorial and Review. Foundations and Trends in Information Retrieval, Volume1, Issue 1, Pages 1-96, 2006.

[5] Roger Kamien. Music: An Appreciation. McGraw-Hill Humanities, 9 editions, 2006.

[6] Alexandros Nanopoulos, Dimitrios Rafailidis, Maria M. Ruxanda, Yannis Manolopoulos: Music search engines: Specifications and challenges. Inf. Process. Manage. 45(3): 392-396, 2009.

[7] J. T. Foote, "An Overview of Audio Information Retrieval." In ACM-Springer Multimedia Systems, vol. 7 no. 1, pp. 2-11, ACM Press/Springer-Verlag, January 1999

[8] Davis Pan, "A Tutorial on MPEG / Audio Compression", IEEE Multimedia, vol. 2, no. 2, pp. 60-74, June 1995.

[9] Michael Fingerhut. Real music libraries in the virtual future: for an integrated view of music and music information. Digitale bibliotheken voor muziek, 2005.

[10] Donald Byrd, Tim Crawford: Problems of music information retrieval in the real world. Inf. Process. Manage. 38(2): 249-272, 2002.

[11] Michael Fingerhut. Music Information Retrieval or how to search for music and do away with incipits. IAML-IASA Congress, Oslo (Norway), August 8–13, 2004.

[12] Aura Lippincott. Issues in Content-based Music Information Retrieval. Journal of Information Science, Vol. 28, No. 2, 137-142, 2002.

[13] Chih-Chin Liu, Po-Jun Tsai, "Content-based Retrieval of MP3 Music Objects," Proceedings of the tenth international conference on Information and knowledge management, October 05-10, 2001.

[14]Michael Clausen, Frank Kurth, Roland Engelbrecht. Context-based Retrieval in MIDI and Audio. In: Fellner, D.W., Fuhr, N., Witten, I. (Eds.): ECDL Workshop: Generalized Documents. Darmstadt, Sept. 2001.

[15] Tseng, Y. H. Content-based retrieval for musiccollections. In Proceedings of ACM SIGIR, 1999.

[16] Remco C. Veltkamp, Frans Wiering, Rainer Typke. Content Based Music Retrieval. In B. Furht (Ed.), Encyclopedia of Multimedia. Springer, 2006.

[17] Yuen-Hsien Tseng: Content-Based Retrieval for Music Collections. SIGIR 1999: 176-182.

[18] Giovanna Neve, Nicola Orio: A Comparison of Melodic Segmentation Techniques for Music Information Retrieval. ECDL 2005: 49-56.

[19] Jia-Lien Hsu, Arbee L.P. Chen, Hung-Chen Chen and Ning-Han Liu. The Effectiveness Study of Various Music Information Retrieval Approaches. Information and Knowledge Management, 2002.

[20] Stephen J. Downie and Michael Nelson. Evaluation of a simple and effective music information retrieval method. In Proceedings of ACM-SIGIR Conference, 2000.

[21] *The Scientific Evaluation* of *Music Information*. *Retrieval Systems*: Foundations and Future. Computer Music Journal, Computer *Music* Journal, 28:2, pp. 12–23, Summer 2004.

[22] Wei-Ta Chu, Wen-Huang Cheng, and Ja-Ling Wu. Semantic context Detection Using Audio Event Fusion. JASP(2006), 2006, pp. 1-12.

[23] Lu, L., You, H., and Zhang, H.-J. A new approach to query by humming in music retrieval. In Proceedings of IEEE International Conference on Multimedia and Expo, 2001.

[24] H. M. Yu, W. H. Tsai, and H. M. Wang, "A query-*by*-Singing system for retrieving karaoke music," IEEE Transactions on Multimedia, vol. 10, no. 8, 2008.

[25] Jeremy Pickens. A Survey of Feature Selection Techniques for Music Information Retrieval, Technical report, 2001.

[26] Mark Levy and Mark Sandler. Music Information Retrieval Using Social Tags and Audio. Trans. Multi. In Multimedia, IEEE Transactions on, Vol. 11, No. 3. (2009), pp. 383-395.

[27] Hwei-Jen Lin, Hung-Hsuan Wu. Efficient geometric measure of music similarity. Information Processing Letters, Volume 109, Issue2, Page 116-120, 2008.

[28] Iman S. H. Suyoto, Alexandra L. Uitdenbogerd, and Falk Scholer. Searching Musical Audio Using Symbolic Queries. IEEE. Transactions on*Audio*, Speech and Language Processing, 16(2):372–381, 2008.