

# Music Indexing and Retrieval: Current Problems

Elizabeth Joan Kelly

J. Edgar & Louise S. Monroe Library

Loyola University New Orleans

## Abstract:

Academics and librarians have yet to reach a consensus on the indexing of print resources about music, nor have they developed satisfactory means of indexing sheet music. With the increasing presence of audio music on the Internet, the need to properly index MP3s and other audio files has reached a new level of urgency, and with it the need to label these items satisfactorily to enable retrieval. While the importance of these fields has been constant since the beginning of indexing and cataloguing, increased availability of sources means that there is more music available to users than ever before, but little in the way of sorting through it. Luckily studies are being undertaken with the aim of solving these problems. This article seeks to explore and explain some of these developments.

## Keywords:

cataloging of music, indexing, digital sound recording, cataloging of sheet music, mp3, information retrieval, music, MIDI

## Article Classification:

Article

This is a post-print [post-peer review] version of a manuscript originally published in *The Indexer* 28, no. 4 (December 2010): 163-166. <http://www.theindexer.org/>

# Music Indexing and Retrieval: Current Problems

## Introduction

The fields of indexing and information retrieval are complex and dynamic as developments in scholarship and technology necessitate constant evaluation and updating of methods and best-practice. Perhaps one of the newest and most challenging areas of indexing and IR is in music. Academics and librarians have yet to reach a consensus on the indexing of print resources about music, such as books on music history, nor have they developed satisfactory means of indexing sheet music. With the increasing presence of audio music on the internet, the need to properly index mp3s and other audio files has reached a new level of urgency and with it the need to label these items satisfactorily to enable retrieval.

There are four primary problems in the indexing and IR of music: indexing of print sources about music, indexing of music-related electronic text, indexing of notated music, and the indexing and retrieval of audio. While the importance of these fields has been constant since the beginning of indexing and cataloguing, increased availability of sources thanks to the internet means that there is more music available to users than ever before, but little in the way of sorting through it. Luckily studies are being undergone with the aim of solving these problems; this paper seeks to explore and explain some of these developments.

## Indexing of Print Sources

One of the first issues in music indexing is the multilingual nature of the source—Western European concert music alone utilizes Italian, French, English, and German. Musical expressions, directions, instruments, and titles can appear in a variety of languages for one single piece. The titles of pieces may vary in language and spelling based on country or region, as can the names of composers. Standard sources, such as the Grove Music Dictionary, need to be consulted when indexing music in

order to avoid confusion (Perry, 1989). By using a well-established source like Grove to determine spelling of music-related words, titles, and names, catalogers and indexers make it possible for sources to be searched for in the OCLC.

Another issue facing music indexers and catalogers is the delicate nature of the primary source. Sheet music can be far more fragile than most books and, because of its size, is often difficult to scan through on a shelf and therefore must be handled frequently. Because of this, proper indexing is necessary for music items so that they can be accessed with a limit of wear and tear. Subject access is also an issue in music queries, as search terms can be comprised of composer, title, instrumentation, performer, format, or level of difficulty. Subject headings can also be used to represent themes and concepts just as they do in fictional works. Dewey considers the following to be the most important facets of music: theory, elements, techniques, character, forms, executants and composer (Myers, 1995). Because it is based on mathematics, music is better suited for computer searching than many other forms; however, this is only useful if search mechanisms are structured properly. Again, a standardized vocabulary of music and terms derived from standard sources, like the *Grove Dictionary of Music*, will help streamline the search process (Myers, 1995).

Yet one more problem in indexing music is the matter of accessibility. Many existing indexing and IR processes are geared towards the trained musician, but what of the novice or casual listener? Searching through a notation-based music index without being able to read music is like an illiterate user trying to use an online catalog. In that situation, the user would probably consult a reference librarian for help in finding what they needed. Since most libraries (outside of some academic ones) will not necessarily staff a trained musician as a librarian, the chances of finding help to search through notated music indexes is far slimmer than in getting assistance with the library catalog. For finding melodies without knowing the title or composer, methods using numerical, letter-based, or directional representations of music for non-musicians are possible (Myers, 1995).

## Indexing of Music-Related Electronic Text

The problem of indexing sources about music is only further complicated when these sources appear online rather than in print. There are three major databases online that index music literature: The Music Index (MI), RILM Abstracts of Music Literature (RILM), and International Index to Music Periodicals (IIMP). There are great differences between the indexing methods of these three databases. In a 2001 study by Jenkins, both qualitative and quantitative studies were performed to analyze the number of articles in each database and how they were indexed, followed by more in-depth characterization of the subject terms applied to the same articles across the three sites. RILM's summarizing abstracts were the most informative for representing the content of the article, while IIMP's descriptive abstracts were better for finding interviews, obituaries, documentary items, and some performance class articles. IIMP uses simple language and keyword searches effectively, making it ideal for the less-experienced researcher. MI lacks abstracts but uses precise subject indexing and a "cross-reference" tool that also are appropriate for beginning researchers. RILM's more scholarly and intricate language is better for advanced researchers, especially musicologists (Jenkins, 2001). Proper designation of each database, and instruction of students who might use them, can enable users to use the best database for their needs and skill sets.

A number of other studies have been done to measure efficient means of multimedia database indexing of music data, but so far these databases have focused on single feature indices, and the few utilizing multi-feature indices are slow and memory consuming. Lo, Lee and Wang (2009) propose a two-tier music index using a combination of Multi-Feature Numeric and Grid-Twin Indexing. Doing this solves the space problem of current systems that use Grid-Twin Suffix trees, and also enhances query searching by eliminating limits on query length. Features including melody, rhythm, chords and contour are extracted to develop the index, meaning that the indexes are searchable by both trained and untrained musicians (Lo, Less, & Wang, 2009).

## Indexing of Notated Music

In addition to indexes of sources about music, studies have been performed on the most effective way of indexing notated music, most often melodies, for easy retrieval. Most of these studies thus far have dealt with the indexing of hymn tunes and folk songs. These sorts of tunes tend to be used many times and appear under different names, so proper indexing allows the user to cross-reference sources rather than aimlessly searching for a melody under a different name.

When a hymn tune is used in multiple songs, there are often subtle changes between versions. Slight variations in a melodic line create the same problem for music indexers that synonyms do for text indexers. After experimenting with both “key of C” conversion methods and Parsons’ contoured method, Coates (1998) created a system using asterisks and letters to quickly index hymn tunes, but the problem of plain versus embellished forms of melodies still persists. The system did prove more accurate than hymn tune name indexes, where nearly 6% of hymn tune names were duplicated (Coates, 1998).

A similar project, the National Tune Index, represents American imprints, ballad operas, dance music, British instrumental music, British songs, and manuscripts. The first stage of the indexing, available on microfiche since 1980, involved the indexing of 38,500 18<sup>th</sup> c. secular works (Camus, 1996). The second phase, Early American Wind and Ceremonial Music 1636-1836, is a computer-generated index of over 20,000 citations taken from over 1,000 works. Pieces are indexed by source, melody, text, and author/performer. Cross references show if part of the piece appears in other works. Tunes are represented by scale degrees of the first three measures of the song, making this index appropriate only for those who can read music. In fact, both Coates’ and Camus’ indexing methods only work if the user can read music. While most users of either of these systems likely can, further research should be done on the possibility of indexing tunes using non-notated means, like contour, rather than note names.

## Indexing and Retrieval of Audio

The majority of current research on music indexing and IR relates to the organizing and cataloging of audio on the internet. The cataloging and classification of music begins with the need for both a physical and intellectual/musical description of the item (Smiraglia, 2002). Music collections are different from any other in the multiple manifestations one may find of one work—several different sized scores published at different times, in different languages, as well as multiple recordings, all for the same piece (Smiraglia, 2002). Enter Music Information Retrieval (MIR), the process of automating IR for musical works and parts of works (Smiraglia, 2002). This includes QBH (Query By Humming, to be discussed further late on in this paper), customized metadata structures, and traditional subject or keyword searching. Traditionally, cataloging and classification have focused on the document, but recent MIR has begun to describe the work instead (Smiraglia, 2002). Items must be catalogued as documentary entities (ie a pocket score of Beethoven's 5<sup>th</sup> Symphony) comprised of both physical and intellectual properties. In order to make these items searchable, we must comprehend what is a musical work and what is an exemplar of a specific musical work, and know how to represent these in the catalog. Uniform titles can be applied to all manifestations of the same work to link them together, creating cross-references similar to those used in Coates' and Camus' notated music indexes. Comprehension of the nature of musical works gains greater importance as we move into the age of digital MIR (Smiraglia, 2002).

While efforts have been made to develop an International digital music representation, currently MIDI is the only widely accepted application (Shakra, Frederico, & El Saddik, 2004). Recently, MPEG-7 tools have been developed to represent data about music using descriptors. The MPEG-7 works by creating algorithms based on different aspects of the music like interval, contour, and repetition (Shakra, Frederico, & El Saddik, 2004). Many indexing techniques exist for music including Short Dynamic Programming and those using Algorithms (Shakra, Frederico, & El Saddik, 2004).

Because of the increasing amount of music available on the internet and the need to make this music searchable, many studies have been done on music indexing, approximation match, and mapping, but few on the actual process of creating algorithms that automatically select the melody of a piece for searching by users (Tang, Lap, and Kao, 2000). Tang et al's study takes a "quick and dirty filtering" approach, favoring efficiency and the ability to process large amounts of music over musical correctness. Five methods are used; three rank tracks according to their similarities, and two classify a track as either melody or non-melody. Tracks are ranked according to velocity (a melody is often the loudest part of the piece), monophonic to polyphonic ratio (tracks with no monophonic moments are probably non-melody), silence to total time ratio, melodic range, and track name (a track called "voice" is probably melodic). The five algorithms were tested on 16012 MIDI files taken from the internet. A random sample of 400 of these were also analyzed by humans to determine the accuracy of the study. Track name was the most accurate algorithm, but should be used with other methods as less than half of the files tested actually had track name information. Velocity was the most accurate of the similarity-based algorithms, but was made less effective by the normalizing of MIDI tracks by computers during sequencing (Tang, Lap, and Kao, 2000).

The increase of multimedia content, particularly music, on the Web means that developments of both the research and marketing level of digitized music need to be made (Shakra, Frederico, & El Saddik, 2004). While there are a number of commercially available audio database management systems, content-based retrieval is a relatively unexplored field (Wieczorkowska & Zytchow, 2003). Because automatic classification cannot be used on raw sound data, sounds must be assigned descriptors (or parameters) based on various factors like acoustics, style, melody, harmonics, instrumentation, and others. This enables content-based searching for audio information. The system detailed by Wieczorkowska and Zytchow uses 49er, a database mining system, to search for regularities among data sets. A collection of CDs featuring chromatic playing of individual instruments was analyzed

for timbral differences. Attributes were then applied to different instrument classes. Finding dependencies in timbral description should help in the development of MPEG-7 for multimedia content description (Wieczorkowska & Zytchow, 2003).

There are also a few attempts to index audio for those who don't read music. The QBH (Query by Humming) system allows users to find a melody by humming it, and the melody is then searched for in a database index developed through corresponding melodic strings. The need for QBH arises from the current trend in using keyword searches to locate multimedia content, a flawed approach if the user remembers the sound of a melody but not the keywords associated with it in a system that negates similarity searching (You, Park, & Kim, 2008). QBH finds melodies similar to the one the user is humming and is designed to factor in user inaccuracies and mistakes so that untrained singers or those with speech impediments can still use the query. The downside of QBH systems, particularly the current favorite N-gram method, is that they require extensive computation resulting in lengthy search times (You, Park, & Kim, 2008). A new method of searching utilizes strings and string matching algorithms instead of numeric representations of melodies, speeding up search times and enabling the development of indexes based on "significant," or frequently searched for, melodies (You, Park, & Kim, 2008). The 3-step search method used in this study proved to be faster and more effective than the N-gram method (You, Park, & Kim, 2008).

Another method for identifying music samples uses the wavelet dispersion vector, an audio content descriptor (Rein and Reisslein, 2006). This study found that the wavelet showed precision in characterizing the audio content as well as good recognition of other performances of the same piece. Testing four different performances of 32 pieces with varying time, frequency, audio environments, and recording quality yielded a 78% mean success rate, showing that this so far underestimated system for music identification would benefit from further research (Rein and Reisslein, 2006).



## Conclusion

While steps are being made towards more efficient means of indexing and retrieving music, there is still a long way to go before best practice is established. The task will only get more difficult as the amount of music and accessibility towards it increases on the internet. By following the steps detailed in the articles summarized in this paper, and by continuing to study the effectiveness of different indexing and IR practices, hopefully optimal means of indexing and searching for notated scores, audio, and writings about music can be developed.

## References

- Camus, R. (1996). Early American wind and ceremonial music, 1636-1836: phase 2 of The National Tune Index [Electronic version]. *Notes*, 52, 723-43.
- Coates, P. (1998). MeloDex: indexing hymn tunes [Electronic version]. *The Indexer*, 21(1), 37.
- Jenkins, M. (2001). A descriptive study of subject indexing and abstracting in International Index to Music Periodicals, RILM Abstracts of Music Literature, and The Music Index Online [Electronic version]. *Notes*, 57(4), 834-63.
- Lo, Y.L., Lee, C.H., & Wang, C.H. (2009). Scalable multi-feature index structure for music databases [Electronic version]. *Information Sciences*, 179(15), 2662-2675. Retrieved from <http://dx.doi.org.proxy.lib.fsu.edu/10.1016/j.ins.2009.03.019>
- Myers, J. (1995). Music: special characteristics for indexing and cataloguing [Electronic version]. *The Indexer*, 19, 269-74. Retrieved from [http://www.theindexer.org/files/19-4/19-4\\_269.pdf](http://www.theindexer.org/files/19-4/19-4_269.pdf)
- Perry, H. (1989). Musical bumps: indexing musical terms [Electronic version]. *The Indexer*, 16, 251-253. Retrieved from [http://www.theindexer.org/files/16-4/16-4\\_251.pdf](http://www.theindexer.org/files/16-4/16-4_251.pdf)
- Rein, S. and Reisslein, M. (2006). Identifying the classical music composition of an unknown performance with wavelet dispersion vector and neural nets [Electronic version]. *Information Sciences* 176 (12), 1629–1655.
- Shakra, I., Frederico, G., & El Saddik, A. (2004). Music Indexing and Retrieval. *IEEE International Conference on Virtual Environments, Human-Computer Interfaces, and Measurement Systems*. Boston, MA. Retrieved from <http://www.mcrlab.uottawa.ca/confs/2004/VE-4008.pdf>

Smiraglia, R P (June 2002). Musical works and information retrieval [Electronic version]. *Notes*, 58(4), 747-764.

Tang, M., Lap, Y.C. and Kao, B. (2000). Selection of melody lines for music databases, *IEEE Multimedia*, 243–248. Retrieved from  
<http://ieeexplore.ieee.org.proxy.lib.fsu.edu/stamp/stamp.jsp?arnumber=00884725>

You, J., Park, S., and Kim, I. (2008). An efficient frequent melody indexing method to improve the performance of query-by-humming systems [Electronic version]. *Journal of Information Science* 2008, 34(6); 777-798.