The Role of Music IR in the New Zealand Digital Library project

David Bainbridge
Department of Computer Science
University of Waikato, Hamilton
New Zealand
d.bainbridge@cs.waikato.ac.nz
Phone: +64 7 838 4407
Fax: +64 7 838 4155

Extended Abstract

This extended abstract describes the computer music work that forms part of the New Zealand Digital Library (NZDL) project. In keeping with the scope of the general project, the music work investigates data acquisition, retrieval, presentation and scalability. These parts are described in turn in the text below.

Introduction

MELDEX—the name given to our digital music library project—is a Web-based system that supports searching through text and sung queries, and browsing through automatically compiled lists of metadata, such as titles. Sample collections include 1,000 popular tunes derived from sheet music using Optical Music Recognition (OMR) software, 10,000 folksongs donated by two other computer music projects, and 100,000 MIDI files gathered from the Web at large.

Matching is monophonic to monophonic. More specifically, it is monophonic query to monophonic track. Collections are permitted to include polyphonic tracks, however no matching against these tracks occurs at present. The software, written in C++ and Perl, implements a distributed architecture making it possible to serve different indexes (text or music) to the same collection on different computers is so desired.

Acquisition

We have worked with three principal acquisition methods: automatic conversion of sheet music using OMR software, on-line MIDI files, and existing databases of music in symbolic form.

For sheet music we selected a book containing 1,200 popular tunes and digitized each page in black and white at 300 dots per inch (dpi). The resulting images took 48 hours to process using a 133 MHz Pentium processor. Analyzing the errors related to notes, 9 edits per 100 notes were necessary to correct durational mistakes, but only 1 edit per 100 notes for pitch. This led us to the conclusion that a collection based on the uncorrected data indexed by pitch would yield a useful collection as the music matching algorithm employed already supports approximate pitch matching.

To boost accuracy rates we have since extended our OMR software to optionally merge the reconstructed score with a MIDI rendition of the same piece of music. The technique developed uses a modified approximate string matching algorithm to align the notes detected in the image with the notes played in the MIDI file and to resolve the differences that arise.

Compared with sheet music, acquiring symbolic music information from a MIDI file is simpler. Pitch calculations are unambiguous, however mapping the timing information to the musical

durations quarter note, half note, *etc*. is more problematic. Example collections formed from MIDI files are: MIDIMINI, based on 1,200 files from a Web site dedicated to MIDI data; and MIDIMAX, based on 100,000 MIDI files located through an Internet search engine.

Working with existing databases of music in symbolic form minimizes the effort in developing a digital music collection as the task of acquisition has essentially been done for us. All that is needed is a utility or two to convert the source database into our internal format. Our 10,000 folk song collection is an example of this type of source data.

Retrieval

Document retrieval is through the activities of searching and browsing. Text searching is accomplished using the full text retrieval system MG with supplementary code to support fielded searching. For melody matching there is a choice of two algorithms: a state-based approach and a dynamic programming approach based on approximate string matching. The latter gives a more precise answer, however it is more computationally expensive. The algorithm used is controlled through a preferences page, as are the parameters use/ignore duration, contour/interval match, adaptive/fixed tuning, and match at start/anywhere.

The main browsing activity in MELDEX takes the form of alphabetically sorted lists of metadata (principally title). Using text-mining techniques we plan to improve on this by extracting additional metadata, cross referencing it and linking it up with other music resource sites. An implemented example of this exists in the MIDIMINI collection where clicking on the MP3 icon next to the title spawns a new search on an MP3 search site using the song title as the query.

Another form of browsing supported in MELDEX encapsulates the notion of "find more tunes that are similar to this one." Based on previously calculated sequences of "interesting notes" that are repeated within a score—a motif as it were—a new query for a given document can be initiated by the clicking on a link next to its title to find tunes similar to the target tune; perhaps even a different version of the same tune. Much hinges on how the pre-calculated motifs are derived (described more fully in the section on scalability below), and this browsing activity is not the only place we can take advantage of this information. We use it to support a form of music document summary for "quick playback" as well as a way of reducing the size of the melody index for searching large collections. We expand on both these points in the text below.

Presentation

Having located a music document, how should it be presented to the user? The answer, of course, depends on the intended use. Three principal formats are audio playback, typeset music notation, and textual information. For a given digital library collection, generation of these formats is strongly dependent on how the music was acquired. For instance, generation of typeset music in an OMR based collection is straightforward, but this is trickier if—as occurs in a MIDI based collection—the source music has never existed in a written form, requiring some form of software tool. In such situations we use an extended command-line version of the music notation application Rosegarden.

To provide an audio summary of a document we draw once again on our pre-calculated motifs, playing, for the segments of the score that correspond to the chosen motifs, all instruments that are involved during that passage of the score.

Scalability

So far we have described the core components of the MELDEX system. In handling large collections, such as 100,000 MIDI files, additional components come in to play to support scalability. These are: the "followed by" operator, bounded memory, and reduction to motifs.

The "followed by" operator works in collections where there is at least one text index. In this scenario the text query is dispatched first, and the document identifiers that match this query are used to form a local database of music documents. The music query is then applied to this smaller, local database.

The bounded memory technique is a rudimentary approach that trades off memory usage for time. Instead of reading in the entire database and matching against that, only part of the database is read in and matching is performed against that. The returned results are stored and then the process is repeated for the next section of the database with the new results being merged with the stored results.

The final approach we have experimented with is to reduce the size of the musical index through the identification of repeated motifs. More specifically we have reduced monophonic tracks to a text string that represents the pitches of the notes played, and then applied a highly efficient text-based phrase detection algorithm. Up to the five longest and five most frequent phrases are chosen, and these are mapped back to their first occurrence in the music file. Applying this technique to the MIDIMINI collection, the index size decreased by 96%.

Conclusion

In conclusion, MELDEX is a rich environment in which to explore the emerging field of digital music libraries, and usage of the service has grown to the point where statistically significant transaction log analysis can be carried out.

Further technical developments are planned. A key step is to develop tighter integration with the main NZDL project, thereby benefiting from a wider base of software developers. This would, for example, increase the number of platforms the software runs on, and adds the option of producing standalone music collections that run off CD-ROM. To learn more about our work visit the Website at `www.nzdl.org`.

Suggested Readings

Bainbridge, D., Nevill-Manning, C., Witten, I., Smith, L. and McNab, J. 1999. Towards a digital library of popular music. *4th ACM conference on Digital Libraries*: 474–478.

Witten, I., McNab, R., Jones, S., Apperley, M., Bainbridge, D. and Cunningham, S. 1999. Managing complexity in a distributed digital library. *IEEE Computer* 42(2): 74–79.