

# Music IR: Past, Present and Future

Alexandra L. Uitdenbogerd, Abhijit Chattaraj, Justin Zobel  
Department of Computer Science, RMIT University  
{alu,achattar,jz}@cs.rmit.edu.au

## Abstract

Music Information Retrieval has a longer history than most people realise, with systems developed in the 1960's. The field has its roots in information retrieval, musicology and music psychology. Information retrieval has provided us with a framework for evaluating retrieval systems. Musicologists have applied various techniques to measure stylistic parameters of composers and general similarity of musical works. Music perception research has taught us that contour is the most important feature of melody for listeners (Dowling 1978). Precursors to computerised music IR are the incipit and theme indexes such as Barlow and Morganstern's dictionary of musical themes (Barlow and Morganstern 1948), however, on-line collections with incipit indexes were soon to follow (for example Hudson 1970).

The current trend in music IR research has been to develop systems that allow the location of answers to queries presented as hummed melodies, or entered in some other way. The earliest system we have found was that developed by Yamamoto in 1988, which only allowed exact matches (Kageyama et al 1993). The nineties have shown a rapid increase in music IR research, with systems papers being published every year since 1993 and more than a dozen published in 1998 (eg. Uitdenbogerd and Zobel 1998, Blackburn and de Roure 1998, Hewlett and Selfridge-Field 1998). In addition, interesting developments have occurred in the field of audio retrieval. Despite this growth of interest in the field, the comparison of systems is difficult because of the lack of a common data set, queries, or relevance judgements.

MIR systems can be evaluated in a similar manner to other systems. As with traditional IR, we need to decide what is a query, what is an answer, and what is meant by relevance, that is, what constitutes similarity. The usual definition of an ad hoc music query is a melody fragment consisting of a sequence of notes. The relevance of answers produced by a system will depend on the user's need. For example, relevant answers for a search for copyright similarity may be different to those for a user who half remembers a song.

In earlier work, we collected queries and relevance judgements from a collection of 10,466 MIDI files by assuming that alternative arrangements or performances of a piece chosen as a query were relevant, and used these to measure effectiveness with the standard techniques of eleven-point precision averages and precision at  $k$  pieces retrieved. These were used to determine which matching techniques were most successful. Whether these relevance judgements correspond to - or yield the same ranking of techniques as - human relevance judgements was an open question.

In recent work, we collected manual queries by having an expert musician listen to MIDI files and create a melodic query for each file. The files were selected so as to allow comparison with our earlier work with automatic queries and relevance assessments. We also developed a system for collecting music relevance judgements from listeners to permit a more established approach to system evaluation. Users were presented with a MIDI file melody and were required to determine whether each answer was highly similar or not.

Marked differences were shown to occur when systems were evaluated with manual as opposed to automatically extracted queries. The differences between the two sets of relevance judgements were more subtle. In general, manual judgements allowed greater discrimination for manual queries and automatic judgements discriminated better between automatic query tests.

Ideally, the future of Music IR research should see the establishment of a common framework for evaluating systems. In terms of research that needs to be pursued, there is a great deal to be done on melody extraction, query and answer presentation, analysis of audio data, as well as the techniques for matching melodies.

### **Suggested Readings**

Barlow, Harold and Morganstern, Sam. 1948. *A Dictionary of Musical Themes*. New York, Crown.

Blackburn, Steven and De Roure, David. 1998. A tool for content-based navigation of music. *Proc. ACM International Multimedia Conference*, Bristol, England.

Dowling W. 1978. Scale and contour: Two components of a theory of memory for melodies. *Computers and the Humanities*, 16: 107-117.

Hewlett, Walter B. and Selfridge-Field, Eleanor (Eds). 1998. Melodic Similarity: Concepts, procedures, and applications. *Computing in Musicology*, 11.

Hudson, Barton. 1970. Toward a Comprehensive French Chanson Catalog. In *Lincoln 1970*.

Kageyama, Tetsuya and Mochizuki, K. and Takashima, Yosuke. 1993. Melody Retrieval with Humming. *Proc. International Computer Music Conference*.

Lincoln, H. (Ed). *The Computer and Music*. 1970. Cornell University Press, Ithaca, New York.

Uitdenbogerd, Alexandra and Zobel, Justin. 1998. Manipulation of Music for Melody Matching. *Proc. ACM International Multimedia Conference*, Bristol, England.

Uitdenbogerd, Alexandra and Zobel, Justin. 1999. Melodic Matching Techniques for Large Music Databases. *Proc. ACM International Multimedia Conference*, Orlando, Florida, USA.