Polyphonic Music Retrieval: The N-Gram Approach

Shyamala Doraisamy
Department of Computing
Imperial College London
shyam_my@yahoo.com

Supervisor: Stefan Rüger
September 2004

This Music Information Retrieval (MIR) study investigates the use of *n*-grams and textual Information Retrieval (IR) approaches for the retrieval and access of polyphonic music data. IR, synonymous with text IR, implies the task of retrieving documents or texts with information content that is relevant to a user's information need.

With music retrieval, the use of *n*-grams has largely been confined to monophonic musical sequences. The few studies that have investigated its use with polyphonic music collections typically reduce a polyphonic file into a monophonic sequence for *n*-gram construction. Techniques for full-music indexing of polyphonic music data with *n*-grams are investigated. A method to obtain *n*-grams from polyphonic music data is introduced. The information content of 'musical *n*-grams' is extended to include rhythmic information in addition to intervallic information. For this, ratios of onset times between two adjacent pairs of pitch events are used. To encode 'musical *n*-grams' to obtain 'musical words' for indexing, a function that maps interval classes to text characters is formulated, and ranges of ratio bins are defined.

These encoding approaches enable encoding of the pitch and rhythm information at various levels of coarseness. Various *n*-gramming strategies are proposed to overcome several problems that arise from the use of the *n*-gram method with polyphonic music. In exploiting the time-dependent element of polyphonic music data, a method to index adjacent and concurrent musical words using a 'polyphonic musical word indexer' is proposed. For the retrieval of these 'overlaying' musical words, i.e., when more than one word can assume the same within-document position, a new proximity-based operator and a ranking function is proposed.

The evaluation results of the indexing approaches proposed are presented, performed on a test collection we developed using approximately 10,000 polyphonic MIDI files. Experiments show that different *n*-gramming strategies and encoding precision differ widely in their effectiveness. The retrieval performances of monophonic and polyphonic queries made to a polyphonic music collection were investigated using text retrieval performance measures. For monophonic queries, we focused in particular on query-by-humming systems, and for polyphonic queries on query-by-example. Error models of these systems were surveyed and included in the fault-tolerance study that investigated the robustness of the *n*-gram method. The feasibility in utilising position information of 'overlaying' musical words was investigated using various proximity-based and structured query operators available with text retrieval systems. Results show that the *n*-gram approach to polyphonic music retrieval is a promising and robust approach for indexing large collections of music.