

Methods for Retrieval and Analysis of Symbolic and Acoustic Music Databases



Ioannis Karydis
Data Engineering Lab.
Department of Informatics
Aristotle University of Thessaloniki



What is a Music Database?

- Music databases are designed to organise musical data.
- Key differences from other DBs:
 - the nature of the data handled
 - the interweave of music with human perception.
- Two research directions in music databases
 - Symbolic data: instructions on what and how to play
 - Acoustic data: digitized audio recording of actual sound waves of a performance.



The Challenges in Music Data Management



- Effectiveness
 - Music data segmentation into meaningful units
 - Polyphony
 - Approximate repetition of occurrences
 - Data quality: music data have performance-related sources of "errors"
- Efficiency
 - Large size of each piece
 - Complex nature of data

Contributions



- Research In Music Databases For Symbolic Data:
 - Mining max length repeating patterns
- Research In Music Databases For Acoustic Data:
 - Similarity Searching
 - P2P Networks

Research In Music Databases For Symbolic Data



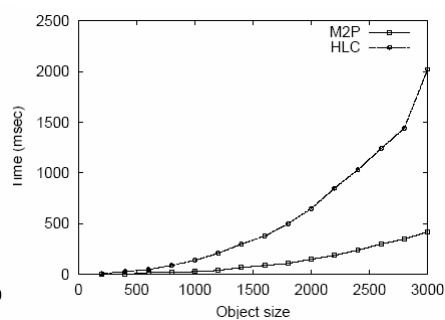
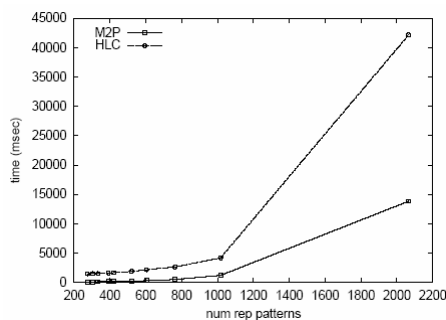
• MLRP

- *Repeating patterns* of a music object
 - Extensive use through out the history of music
 - Reference point for the discovery of music *themes*
 - Have been considered as *characteristic signatures* of music objects
- Introduced of the problem of finding the maximum-length repeating patterns (MLRPs)
- Presented an efficient, novel algorithm, M2P, for the extraction of MLRPs.
- M2P avoids costly repetition of frequency calculations by examining as few as possible intermediate repeating patterns, and aiming at fast reaching of MLRP set.

Research In Music Databases For Symbolic Data



• MLRP



Research In Music Databases For Symbolic Data



● Repeating Pattern Applications

- Future research direction
- Repeating Patterns (RPs) used for indexing purposes
- Motivation:
 - their adherence to both efficiency and semantic requirements for CBMIR,
 - RPs correlate to *motifs*,
 - The total size of the collection of repeating patterns is smaller than that of the music objects,
 - Real world scenario: high probability that user queries are bias towards motifs or themes

Research In Music Databases For Acoustic Data



● Similarity Searching

- Humming MIR is the most natural and spontaneous CBMIR,
- CBMIR can also be performed by providing a query music file or even by creating a query using a keyboard.
- Freedom of origin = need for similarity searching
 - due to possible errors in the query.
- Songs are developed based on variation of an original theme.

Research In Music Databases For Acoustic Data



● Similarity Searching

- Feature extraction method for acoustical music sequences
- Extracted features are grouped by Minimum Bounding Rectangles (MBRs)
- Indexed using a spatial access method
- Presented a novel false alarm resolution method that utilises a reverse order schema

Research In Music Databases For Acoustic Data



● P2P Networks

- Wired
 - Presented a novel framework for efficient CBMIR in unstructured P2P networks
 - Minimisation of the required traffic for all operations with the use of an intelligent sampling scheme
 - No false negative results occur
 - Analysed the relative performance of the various similarity searching algorithms in the framework.

Research In Music Databases For Acoustic Data



- **P2P Networks**

- Wireless
 - The wireless medium = new requirements
 - CBMIR methods for wired P2P networks do not consider the continuous alteration of the network topology,
 - Need to reduce traffic by replacing the original query with a new representation that utilises novel, appropriate transcoding schemes.
 - In wireless ad-hoc networks answers to the query have to be propagated back to the querier via the network

Research In Music Databases For Acoustic Data



- **P2P Networks**

- Wireless
 - Introduced CBMIR application in wireless ad-hoc networks
 - Proposed a novel algorithm, with twofold optimisation:
 - the use of query representation with reducing length,
 - a selective policy for the routing of answers, which performs additional pruning of traffic.
 - Result:
 - significant reduction in response times and traffic
 - The examined context does not depend on specific features and distance measure

Perspectives



- Interesting/promising research directions:
 - research of the structure of acoustic data (further than thumbnails)
 - accurate music transcription (acoustic to symbolic)
 - methods for similarity searching in polyphonic-symbolic data.

**Methods for Retrieval and Analysis of
Symbolic and Acoustic Music Databases**



Thank you!