

MODULO7 : A FULL STACK MUSIC INFORMATION RETRIEVAL AND STRUCTURED QUERYING ENGINE

First Author

Affiliation1

author@ismir.edu

Second Author

Retain these fake authors in

submission to preserve the formatting

Third Author

Affiliation3

author3@ismir.edu

ABSTRACT

This paper describes a novel Music Information Retrieval and Structured Querying framework named Modulo7. Modulo7 is a full stack implementation (both client and server side software) which facilitates indexing variegated sources of music (midi, mp3, music xml and digitized sheet music files). Modulo7 implements a similarity search engine based on customized vector space representations of songs, an efficient indexing and persistent storage mechanism and an interface for querying attributes based on SQL(Structured Querying Language) like principles. The papers describes the implementation details and outlines speed up and scale up results over input sources and other MIR frameworks.

Keywords : MIDI, Music XML, MP3, Music Retrieval, SQL

1. INTRODUCTION

Given the explosive growth of Music Information Retrieval research, several approaches and software suits have been designed to tackle generic problems such as efficient indexing, similarity searches, archival methods, structured and un-structured querying, feature extraction, audio and digital signal processing. A vast majority of the MIR frameworks in academia tend to approach very specific problems and does not support scalability as a significant end goal in itself [4]. Moreover, industry applications predominantly treat MIR applications based on collaborative filtering approaches [8] and/or manual annotation [5] instead of structural analysis of music sources yet scales very well to large datasets.

Modulo7 is an attempt to capture the best features of both worlds. Modulo7 converts different music sources (midi, musicxml files, sheet music png/jpeg files and mp3) into a single unified symbolic representation. It indexes songs on global properties (artist, key signature, time signature etc) and provides a vector space implementation for

similarity searches and querying based on the simlie [7] platform and extended to include chords and polyphony. As a consequence, Modulo7 acts as an end to end software suite for Music Information Retrieval including client side querying, searching, indexing and persistent store mechanism.

2. RELEVANT WORK

Music Information Retrieval is a vast and interdisciplinary body of work. In order to facilitate research in MIR, several software suits and frameworks have been developed in the academic community in the past. Notable amongst them are software suits like jMIR [10] for automatic feature extraction from audio and midi sources to be used as input into machine learning algorithms for genre classification, marsyas [13] and essentia [1] for audio processing, humdrum [2] for automated musicological research, gamera [6] for optical music recognition and symbol identification and SIMILIE [7] for melodic similarity analysis.

On top of the effort done by there has been significant effort undertaken by companies in the field of MIR with an emphasis on music recommendation. For instance Amazon implements a collaborative filtering approach by quantifying a consumer's shopping behavior [8]. Pandora on the other hand utilizes an approach in which trained musicologists manually classify similarity between artists and songs [?]

3. SOFTWARE ASPECTS

This section details the software aspects of the Modulo7 most notably its architecture and a typical work-flow description.

3.1 Software Architecture and Design

The Modulo7 architecture can be visualized as 1 and is broadly divided into the following components.

1. **Music Parsers:** Components that individually parse different music sources into a unified symbolic representation. as described in 2. This object can be stored in memory or in a persistent serialized state via Apache Avro. Depending on the source of the file, the parser utilizes a different approach.



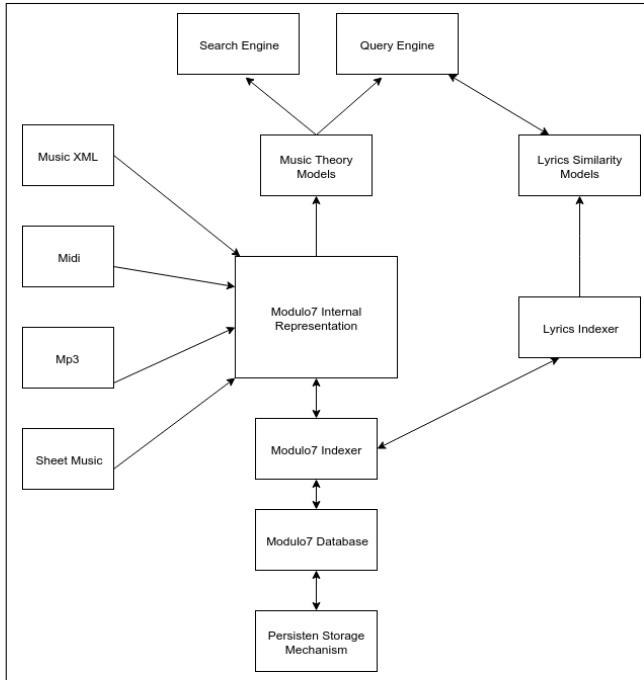


Figure 1. A block diagram of the Modulo7 Architecture.

- (a) For mp3 files, the input is parsed into chromagrams using echonest API and then converted into symbolic format using the KKTonality profiles algorithm [9].
 - (b) For Midi and MusicXML files, an in house parser is built.
 - (c) For Sheet Music, the Audiveris library is used to convert digitized sheet music to
2. **Music theory Models:** Modulo7 implements vector space models described in [7] with extensions to polyphonic music and chords using a template matching algorithm described in [12]. This forms a mathematical formulation of music sources described in detail in section 4 which is leveraged for querying and similarity searches.
 3. **Client side Querying:** Modulo7 implements client side abstractions that exposes two modes of information retrieval.
 - (a) **Structured Querying:** The structured querying mode allows a user to make SQL like queries based on certain criteria or statistic that is quantified on music theoretic criteria.
 - (b) **Custom Search Engine:** Modulo7 exposes a search engine functionality but allowing for customized similarity metric choices as described in section 5.1
 4. **Lyrics Engine:** On top of providing support for music sources, Modulo7 also implements a standard text based retrieval model on lyrics of songs written in Apache Lucene.

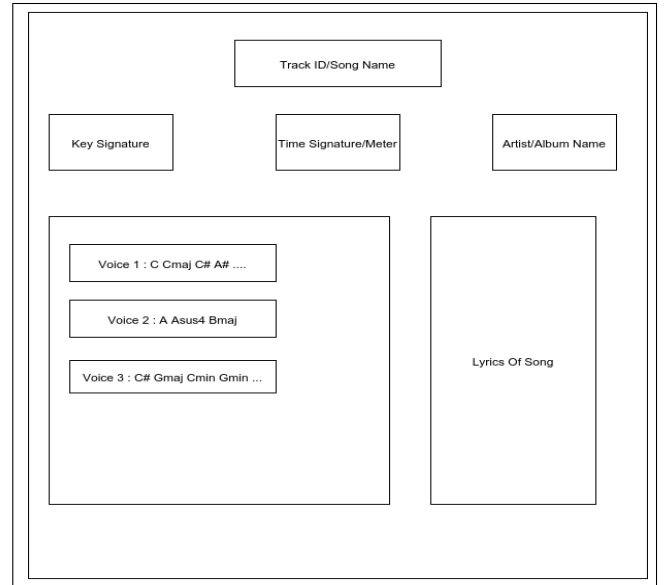


Figure 2. Modulo7 internal representation.

3.2 Workflow

The workflow can be defined as the steps of operations the user takes to initialize the Modulo7 system and expose the consumer to client facing querying options

1. Modulo7 is initialized by specifying a source directory. Every music source file inside the directory is parsed and converted to either an in memory or persistent store version

4. VECTOR SPACE MODELS FOR MUSIC SOURCES

5. CLIENT ABSTRACTIONS

5.1 Search Engine Functionality

6. EXPERIMENTAL EVALUATION

This section includes improvements that have been attained over existing frameworks and input sources. For the purposes of this experiment, existing datasets such as the Saarland dataset [11], the Wikifonia Dataset [14] and the million song dataset [3].

6.1 Indexing and Persistent Store Improvements

7. CONCLUSION AND FUTURE WORK

8. REFERENCES

- [1] *ESSENTIA: an Audio Analysis Library for Music Information Retrieval.*
- [2] Anon. The humdrum toolkit: Reference manual. menlo park, california: Center for computer assisted research in the humanities, 552 pages, isbn 0-936943-10-6.
- [3] Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The million song dataset. In *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011)*, 2011.

- [4] Donald Byrd and Tim Crawford. Problems of music information retrieval in the real world. *Inf. Process. Manage.*, 38(2):249–272, March 2002.
- [5] W.T. Glaser, T.B. Westergren, J.P. Stearns, and J.M. Kraft. Consumer item matching method and system, February 21 2006. US Patent 7,003,515.
- [6] Ichiro Fuginaga Karl MacMillan, Micheal Droettbroom. Gamera: Optical music recognition in a new shell.
- [7] Daniel Mllensiefen Klaus Frieler. The simile algorithms for melodic similarity.
- [8] Greg Linden, Brent Smith, and Jeremy York. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1):76–80, January 2003.
- [9] Sren Tjagvad Madsen, Gerhard Widmer, and Johannes Kepler. Key-finding with interval profiles.
- [10] Cory Mckay. Automatic music classification with jmir, 2010.
- [11] Meinard Müller, Verena Konz, Wolfgang Bogler, and Vlora Arifi-Müller. Saarland music data (SMD). In *Late-Breaking and Demo Session of the 12th International Conference on Music Information Retrieval (ISMIR)*, Miami, USA, 2011.
- [12] Adam M. Stark and Mark D. Plumbley. Real-time chord recognition for live performance. 2009.
- [13] George Tzanetakis and Perry Cook. Marsyas: A framework for audio analysis. *Org. Sound*, 4(3):169–175, December 1999.
- [14] Wikifonia. The wikifonia lead sheet collection, 2013.