

A Robust Retrieval System of Polyphonic Music Based on Comparison of Chord Sequences

Pierre Hanna, Thomas Rocher and Matthias Robine
LaBRI - Université de Bordeaux 1
F-33405 Talence cedex, France
firstname.name@labri.fr

ABSTRACT

Retrieval systems for polyphonic music rely on the automatic estimation of the similarity between two musical pieces. In the case of symbolic music, existing systems consider a monophonic reduction based on melody or propose algorithms with high complexity. In this paper, a new approach is presented. Musical pieces are represented as a sequence of chords estimated from groups of notes sounding at the same time. A root and a mode is associated to each chord. Local alignment is then applied for estimating a similarity score between these sequences. Experiments with MIDI files collected on the Internet show that the system proposed allows the retrieval of different versions of the same song.

1. INTRODUCTION

The number of digital musical pieces available on the Internet is always growing. New interfaces for browsing and recommending music in large databases have to be developed for users, since retrieval systems based on textual queries impose a prior knowledge of the properties of musical pieces. Researches about Music Information Retrieval involve proposals of new retrieval systems [9]. One of the major problem is the automatic estimation of the similarity between musical pieces. Since research works generally concern Western music, four main characteristics of the music are considered: rhythm, timbre, melody and harmony. In this paper we focus on symbolic music (music represented as musical events), and on tonal properties.

Existing systems for estimating the symbolic melodic similarity generally deals with monophonic music [8, 12]. Adaptations of geometric [7] or string-matching algorithms [3] to polyphonic music have been recently proposed. Some systems consider a monophonic reduction of the musical piece compared. Others propose adaptations that increase the complexity of algorithms [10, 4] and thus prevent their application to large databases. For example, a way would be to consider all the distinct monophonic lines induced by polyphony. But this naive method may imply a combinatorial explosion in the cases of large scores [5].

Therefore, it may be interesting to investigate an adapted representation of groups of notes that sound at the same

time, in order to avoid the complete ignoring of some notes. Like existing systems, we propose a monophonic reduction, but the main difference is that representation based on chords considers more tonal information than when considering only the notes of melody. The contribution of this paper is a robust retrieval system for symbolic polyphonic music, presented in Section 2, and experiments showing its accuracy in Section 3.

2. POLYPHONIC MUSIC RETRIEVAL

The first step of the retrieval system proposed consists of analyzing polyphonic music for representing it as a sequence of chords. Then local alignment algorithm is applied for comparing the chord sequences computed.

2.1 Chord analysis

The system used for chord analysis provides the root of the chord as well as its mode (major or minor). It first defines groups of notes, where each group is formed by notes starting and ending at the same time. It then builds a graph of all possible chords for each group of notes, and selects the best path in this graph. A rule-based approach is adopted to enumerate chord candidates from groups of notes by considering compatible chords and compatible keys. The distance proposed by Lerdahl [6] is then used to compute costs between different chord candidates. Dynamic programming is also involved to select the best path among chord candidates.

Experiments have been performed on a MIDI song database with a given ground truth for chords. This database is composed of 155 songs, belonging to different musical styles. For each MIDI tick, we compare the estimated chord to the ground truth. Each tick is then correctly estimated if the root and the mode match the ground truth, and incorrectly estimated otherwise. The global accuracy on this database is 83.7%. An error analysis is proposed in Table 1. Results of this analysis show that two thirds of errors concern fifth and third minor intervals (respectively 26.5% and 42.9%). The comparison algorithm proposed has to take into account this characteristic.

	Types of Errors					
	1	2	3	4	5	6
Errors (semitones)	1	2	3	4	5	6
%	5.5	9.8	42.9	13.9	26.5	1.3

Table 1: Error occurrences (in %) regarding the root of the estimated chord. When the estimated root does not match the ground truth, a distance is calculated between the estimated root and the correct root. This distance is in semitones.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGIR '09

Copyright 200X ACM X-XXXXX-XX-X/XX/XX ...\$5.00.

2.2 Similarity Between Chord Sequences

Among several existing methods, Smith and Waterman's approach [11] consists in detecting local similar areas between two sequences. A similarity score is calculated by considering elementary operations transforming one string into the other. The operations between sequences include deletion, insertion of a symbol, and substitution of a symbol by another. Algorithms based on local alignment have been successfully adapted for melodic similarity comparison purposes [8]. A first approach for audio chord sequences comparison has also been presented in [1]. Regarding the errors due to the chord analysis method, we propose to compute substitution scores according to the consonance of the interval between the roots of the chords compared. These scores also take into account the difference between modes.

3. EXPERIMENTS

Experiments were performed using 1000 symbolically encoded musical pieces. 850 MIDI files were collected on the Internet and checked for quality. 150 other MIDI files represent versions of 43 different songs, forming 43 classes of songs. Versions can differ in a number of ways: differently sequenced or covers of the same song. Exact copies of files have been deleted from the database by first comparing files using the *diff* Linux command, but also by applying the retrieval system proposed: if the similarity score between two musical pieces is estimated as maximum (i.e. the similarity score of one of the piece compared to itself), one of the musical pieces was deleted. List of MIDI files and detailed results are available at <http://dept-info.labri.fr/~hanna/SIGIR>.

Table 2 shows the results of the experiments performed by considering each version as a query. The total number of tests is 150. For each one, the corresponding version(s) is(are) expected to be retrieved at the first rank(s). Columns of Table 2 show the average first tier and second tier per class, recall at TOP 10 and TOP 20. The first tier is the number of correctly retrieved songs within the best ($C - 1$) matches divided by $(C - 1)$, where C is the number of versions of the same song [2]. The second tier is the number of correctly retrieved songs within the best $(2C - 1)$ matches divided by $(2C - 1)$. Recall at TOP 10 is the number of correctly retrieved songs within the best 10 matches, divided by $(C - 1)$.

The average first tier is 0.68 and indicates that the retrieval system is quite accurate. Songs of 19 classes allow the retrieval of every corresponding song at the first best matches. One third of the query songs (51 over 150) gives all the corresponding versions at the first best matches. More than 75% (116 over 150) queries allow the retrieval of at least one corresponding version at the first rank. The average Mean Reciprocal Rank of the first correct piece is 0.80. Errors occur when musical differences between the song compared are important. For example, we tested our system with two versions of the Coldplay song *Viva la Vida* and the Satriani song *If I Could Fly* which has been estimated as very similar by many listeners. Our system correctly estimates the two versions of Coldplay song as very similar, but different from the Satriani song, because the differences are complex (different musical structures, different verses, etc.).

1st Tier	2nd Tier	TOP 10	TOP 20
0.68	0.73	0.75	0.78

Table 2: Results of experiments for retrieving song versions.

4. CONCLUSIONS AND FUTURE WORKS

In this paper, we propose a robust retrieval system for symbolic polyphonic music. Whereas existing systems mainly rely on monophonic reduction by analyzing the main melody, the system presented estimates a root and a mode for each group of notes sounding at the same time. This approach is less discriminant than the existing methods since two songs may be very different with a similar chord progression. But it may be interesting to combine it with an approach based on melody extraction and comparison. Future work concern adaptations of such approach to polyphonic audio retrieval by representing sound samples as a sequence of pitch class profiles. Limitations of existing methods for audio chord estimation have to be precisely studied in order to adapt algorithms for comparison.

5. REFERENCES

- [1] J. Bello. Audio-based Cover Song Retrieval using Approximate Chord Sequences: Testing Shifts, Gaps, Swaps and Beats. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 239–244, Vienna, Austria, 2007.
- [2] B. De Haas, R. Veltkamp, and F. Wiering. Tonal Pitch Step Distance: A Similarity Measure for Chord Progressions. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 51–56, 2008.
- [3] S. Doraisamy and S. Rüger. Robust Polyphonic Music Retrieval with N-grams. *Journal of Intelligent Information Systems*, 21(1):53–70, 2003.
- [4] P. Hanna, M. Robine, P. Ferraro, and J. Allali. Improvements of Alignment Algorithms for Polyphonic Music Retrieval. In *Proceedings of the Int. Computer Music Modeling and Retrieval Conference (CMMR)*, pages 244–251, Copenhagen, Denmark, 2008.
- [5] K. Lemström and A. Pienimäki. Approaches for Content-Based Retrieval of Symbolically Encoded Polyphonic Music. In *Proceedings of the International Conference on Music Perception and Cognition (ICMPC)*, pages 1028–1035, Bologna, Italy, 2006.
- [6] F. Lerdahl. *Tonal Pitch Space*. Oxford University Press, 2001.
- [7] A. Lubiw and L. Tanur. Pattern Matching in Polyphonic Music as a Weighted Geometric Translation Problem. In *Proceedings of the fifth International Conference on Music Information Retrieval (ISMIR)*, pages 289–296, Barcelona, Spain, 2004.
- [8] M. Mongeau and D. Sankoff. Comparison of Musical Sequences. *Computers and the Humanities*, 24(3):161–175, 1990.
- [9] N. Orio. Music Retrieval: A Tutorial and Review. *Foundations and Trends in Information Retrieval*, 1(1):1–90, 2006.
- [10] B. Pardo and M. Sanghi. Polyphonic Musical Sequence Alignment for Database Search. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR)*, pages 215–222, London, UK, 2005.
- [11] T. Smith and M. Waterman. Identification of Common Molecular Subsequences. *Journal of Molecular Biology*, 147:195–197, 1981.
- [12] A. L. Uitdenbogerd and J. Zobel. Manipulation of Music for Melody Matching. In *Proceedings of the Sixth ACM International Conference on Multimedia*, pages 235–240, 1998.